ABI PRISM®
SeqScape® Software
Version 2.1

User Guide

Applied Biosystems

© Copyright 2003, Applied Biosystems. All rights reserved.

For Research Use Only. Not for use in diagnostic procedures.

Information in this document is subject to change without notice. Applied Biosystems assumes no responsibility for any errors that may appear in this document. This document is believed to be complete and accurate at the time of publication. In no event shall Applied Biosystems be liable for incidental, special, multiple, or consequential damages in connection with or arising from the use of this document.

SeqScape software has not undergone specific validation for human identification applications. Human identification laboratories must perform their own validation studies.

Notice to Purchaser: License Disclaimer

Purchase of this software product alone does not imply any license under any process, instrument or other apparatus, system, composition, reagent or kit rights under patent claims owned or otherwise controlled by Applera Corporation, either expressly, or by estoppel.

TRADEMARKS:

ABI PRISM, Applied Biosystems, BigDye, MicroSeq, SeqScape, and ViroSeq are registered trademarks of Applera Corporation or its subsidiaries in the U.S. and/or certain other countries.

AB (Design), Applera, GeneMapper, POP-4, POP-5, POP-6, and POP-7 are trademarks of Applera Corporation or its subsidiaries in the U.S. and/or certain other countries.

Windows, Windows NT, Windows 2000, and Windows XP are registered trademarks of Microsoft Corporation.

All other trademarks are the sole property of their respective owners.

Part Number 4346367 Rev B

09/2003

Contents

Preface

How to Use This Guide xv
Conventions Used in This Guide xvi
Workstation Safety xviii
How to Obtain More Information xviii
How to Obtain Services and Support xix

Chapter 1 Introduction to ABI PRISM SeqScape Software

New Features in SeqScape Software v2.1 1-2
New Features for v2.1 1-2
For Users of Earlier Versions of SeqScape Software 1-3
Features of v2.0 1-3
About SeqScape Software 1-5
Genetic Analyzer Applications 1-5
SeqScape Software Applications 1-5
Resequencing Data with SeqScape Software 1-5
Data Sources for Resequencing Projects 1-6
Levels of Automated Analysis 1-6
What the Software Does 1-6
How the Software Performs Analyses 1-7

Chapter 2 Getting Started

Administrator: Registering the Software 2-2
License and Warranty 2-2
Registering Your Software 2-2
Hardware and Software Requirements 2-3
Minimum System Requirements 2-3
Hard Drive Partitions 2-4
SeqScape Software Versions 2-5

Preparing to Install SeqScape Software 2-6
Installing the Full Version for the First Time 2-7
Upgrading from SeqScape Software Versions 1.0, 1.1, or 2.0 2-8
About the Upgrade 2-8
Upgrading to v2.1 2-8
Existing Users 2-8
Removing SeqScape Software 2-9
What the Uninstallation Process Does 2-9
Removing the Software 2-9
Starting SeqScape Software for the First Time 2-10
Before You Begin 2-10
File-Naming Convention 2-10
Starting SeqScape Software 2-11
Creating New Users 2-13
Setting Up Authentication & Audit 2-15
Changing User Information 2-19
Setting Up the Default Directory 2-20
New Users Logging In for the First Time 2-21
When New Users Log In 2-21
SeqScape Software Structure 2-22
SeqScape Manager Window 2-22
Project Window 2-23
SeqScape Software Toolbar 2-24
Menus on the Main SeqScape Window 2-26
Workflow 2-28

Chapter 3 Creating Analysis Defaults and Display Settings

Workflow for This Chapter 3-2
Analysis Defaults Settings 3-3
Creating Analysis Protocols 3-3
Analysis Protocol Editor Tabs 3-4
Specifying the Basecall Settings 3-4
Specifying the Mixed Bases Settings 3-7
Specifying Clear Range 3-8
Specifying the Filter Settings 3-10

Specifying the Analysis Settings 3-12
Gap and Extension Penalties 3-12
Setting Analysis Defaults 3-14
Selecting the Analysis Default Settings for Individual Samples . . 3-17
Specifying Display Settings 3-18

Chapter 4 Creating a Reference Data Group

Workflow for This Chapter 4-2
Reference Data Group (RDG) 4-3
About the Reference Data Group 4-3
GenBank Features 4-4
Downloading a GenBank File 4-5
About Creating a New Reference Data Group (RDG)s 4-5
Creating a New RDG Using the Wizard 4-6
Using the Wizard to Learn the Software 4-6
Setting Up the Reference Segment 4-10
Creating a New RDG Using SeqScape Manager 4-12
Before You Begin 4-12
Creating an RDG from SeqScape Manager 4-12
About the Reference Sequence 4-12
Importing a Reference Segment 4-13
Defining Regions of Interest (ROI) 4-15
Defining an ROI 4-15
Pasting a Reference Segment 4-15
Deleting an ROI or Layer 4-16
Deleting a Reference Segment 4-16
ROI Tab Descriptions 4-18
Layer Pane Functions 4-18
The ROI Pane 4-18
Columns in the ROI Pane 4-19
Creating a Library 4-20
About the Library 4-20
Using Aligned FASTA Files 4-20
Using a Tool to Align the Files 4-20
Setting Up Your Library 4-21
Creating New Layers 4-24
Adding a Reference Break in a Sequence 4-27

Declaring Variants into an RDG 4-29
About NT Variants 4-29
Creating New NT Variants 4-30
Importing NT Variants in Tab-Delimited Format 4-32
Creating an RDG from Aligned Consensus Sequences 4-34
About Creating an RDG 4-34
Importing NT Variants from an Aligned FASTA File 4-34
Entering New AA Variants 4-36
Importing AA Variants 4-38
Assigning Styles to Variants 4-39
Saving a Copy of the RDG 4-41
Saving the RDG for Other Projects 4-41
Save To Manager As Button 4-42

Chapter 5 Creating a Project Template

Workflow for This Chapter 5-2
Creating a Project Template 5-3
About Creating a New Project Template 5-3
Creating a New Project Template 5-4
Saving Project Components 5-5
About Saving Template Components 5-5
Saving Template Components from Within a Project 5-5
Examples of Changing the Settings Within a Project 5-6

Chapter 6 Creating and Analyzing a Project

Workflow for This Chapter 6-2
Before You Begin Creating a Project 6-3
What an Analysis Entails 6-3
Ways to Create and Analyze a New Project 6-4
Using the New Project Wizard to Create and Analyze a Project 6-5
The New Project Wizard 6-5
Creating and Analyzing a New Project Using a Project Template . . . 6-10
About the Project Template 6-10
Creating a New Project Using a Template 6-10

Adding Specimens and Importing Data into a Project 6-11
Overview 6-11
Adding Specimens and Importing Samples Automatically 6-11
Adding Specimens and Importing Samples Manually 6-14
Adding Specimens and Importing Data Files 6-15
Importing Text-Only Files 6-20
Removing Samples or Specimens 6-21
Analyzing the Data 6-21
Running an Analysis 6-21
Reanalyzing a Project Using a Different Project Template 6-22
When You Would Want to Do This 6-22
Saving a Project Before Reanalyzing 6-22
Applying a Template to an Existing Project 6-23
Incorporating Variants into the Project RDG 6-25
About Incorporating Variant Sequences 6-25
Changing a Single Unknown Variant to a Known Variant 6-25
Changing Multiple Unknown Variants 6-27
Importing Variants 6-28
Creating a New Variant in a Project 6-29
Adding a Variant in the Project 6-31
Importing Variants to the Project 6-32
Importing and Exporting Project Information 6-34
About Importing and Exporting 6-34
Importing from SeqScape Manager 6-34
Exporting from SeqScape Manager 6-34

Chapter 7 Viewing the Results

Workflow for This Chapter 7-2
About the Project Data 7-3
View Formats 7-3
Data Display Conventions 7-3
Quality Value Display 7-3
Exporting and Printing Project Data 7-3
Project Views 7-4
Displaying the Project Views 7-4
Specimen Views 7-10
Segment Views 7-11
Displaying the Segment Views 7-11

Sample Views 7-15
Displaying the Sample Views 7-15
Viewing Variants 7-20
Saving Your Data 7-21
About the Reports 7-22
Types of Reports 7-22
Exporting and Printing Reports 7-22
Viewing the Reports 7-23
Viewing the Reports and Project Results 7-25
Analysis QC Report 7-26
Mutations Report 7-28
AA Variants Report 7-31
Specimen Statistics 7-32
Sequence Confirmation Report 7-33
Base Frequency Report 7-34
Library Search Report 7-35
RDG Report 7-37
Audit Trail Report 7-38
Genotyping Report 7-39
Customizing the Reports 7-42
Customizing Text Settings 7-42
Customizing the Data View 7-44

Chapter 8 Reanalyzing and Editing Data

Workflow for This Chapter 8-2
About Analysis Parameters 8-3
Introduction 8-3
Viewing Analysis Parameters in the Sample Manager 8-4
Changing the Analysis Parameters in the Sample Manager 8-6
Adding Samples to the Sample Manager 8-6
Changing Basecaller and DyeSet/Primer Files 8-6
Changing the Analysis Parameters in an Analysis Protocol 8-7
Editing an Analysis Protocol 8-7
Applying the Analysis Protocol 8-12

Editing the Data 8-13
About Sequence Editing 8-13
When to Edit the Data 8-13
Editing a Sample or a Consensus Sequence 8-14
Editing a Consensus Sequence in the Segment View 8-14
Editing Sample Bases 8-14
Editing a Consensus Sequence in the Project View 8-15
Adjusting the Clear Range 8-16
About the Clear Range 8-16
Using the Clear Range Widget 8-17
Using the Mouse 8-18
Using the Set Clear Range Dialog Box 8-19
Editing Variants 8-20
Method 1 8-20
Method 2 8-21
Saving Your Data 8-21
Adding a Genotype 8-22

Chapter 9 Exporting and Printing Data and Reports

Exporting Data Files 9-3
File Names 9-3
Format Options 9-3
Exporting a Project Alignment 9-4
Exporting a Specimen 9-5
Exporting a Segment 9-6
Exporting a Sample 9-7
Exporting Reports 9-8
File Names 9-8
Format Options 9-8
Exporting a Report 9-9
Exporting All Reports Automatically 9-10
Printing Data and Reports 9-11
Customizing Header and Footer Display 9-11
Printing Views of a Sample File 9-12
Printing Various Views of a Project 9-14
Printing a Report 9-16

Chapter 10 Sample and Consensus Quality Values

Types of Quality Values (QVs) 10-2
Sample Quality Values 10-3
Sample Quality Values 10-3
Interpreting the Sample Quality Values 10-3
Sample Score 10-3
Consensus Quality Values 10-5
Interpreting the Consensus Quality Values 10-5
Consensus Score 10-5
Displaying Quality Values 10-6
Customizing the Quality Value Display 10-7
Displaying the Quality Bars and Values 10-9
Editing Bases with Quality Values 10-10
Cumulative Quality Value Scoring in Reports 10-11
Analysis QC Report 10-11
Mutations Report 10-12
Specimen Statistics Report 10-13
Genotyping Report 10-14

Chapter 11 Automating Analysis

Integrating SeqScape and Data Collection Software 11-2
Overview 11-2
Software Relationships 11-3
Before You Start 11-4
Required Files Created in SeqScape Software 11-6
Launching Autoanalysis Manger 11-7
Creating Required Files in the Data Collection Software 11-8
For More Information 11-8
If the Files Already Exist 11-8
Creating an Instrument Protocol 11-9
Creating an Analysis Protocol 11-11
Creating a Results Group 11-14
Creating a Plate Record 11-18
Completing a Plate Record 11-20
Scheduling and Starting a Run 11-22

Understanding the Autoanalysis Manager 11-24
Overview 11-24
Files Created 11-24
Components 11-24
General Tab 11-25
SeqScape 2.1 Tab 11-26
File Sharing Between Data Collection and SeqScape Software 11-28

Appendix A Basecallers and DyeSet/Primer Files

Definitions and Naming A-2
Basecaller Algorithm A-2
DyeSet/Primer File A-2
DyeSet/Primer File-Naming Conventions A-2
Basecaller and DyeSet/Primer Compatibility A-4
ABI PRISM 310 Genetic Analyzer Files A-5
ABI PRISM 377 DNA Sequencer Files A-8
ABI PRISM 3100 Genetic Analyzer Files A-10
ABI PRISM 3100-*Avant* Genetic Analyzer Files A-13
ABI PRISM 3700 DNA Analyzer Files A-15
Applied Biosystems 3730/3730xl DNA Analyzers Files A-17

Appendix B Frequently Asked Questions

Upgrading FAQs B-2
Training and Documentation FAQs B-3
SeqScape Software Basics FAQs B-4
General SeqScape Software FAQs B-5
SeqScape Manager FAQs B-9
Library FAQs B-14
Mutation, Variant, HIM, and HFM Detection FAQs B-15
Data Analysis FAQs B-16
Analysis Reports FAQs B-18
Quality Values FAQs B-19
Printing and Exporting Results FAQs B-20
Audit Trail, Security, and Access Control FAQs B-21

Appendix C Translation Tables

IUPAC/IUB Codes C-2
IUPAC Diagrams C-3
Complements C-3
Universal Genetic Code C-4
Amino Acid Abbreviations C-5

Appendix D User Privileges

Tables of User Privileges D-1

Appendix E Aligned Variant and FASTA File Format

Tab-Delimited Files E-2
Creating a Variant Text File E-2
FASTA File Format E-3
FASTA Format Description E-3
FASTA Format Example E-3
IUB/IUPAC Codes E-4
Supported Nucleic Acid Codes E-4
Supported Amino Acid Codes E-5

Appendix F Creating a Multi-Aligned FASTA File

Introduction F-1
Method 1: Create a Multi-Aligned FASTA File Using SeqScape Software . . F-2
1. Create a Multi-Aligned FASTA File F-2
2. Create Your Library in SeqScape Software F-3
Method 2: Create a Multi-Aligned FASTA File Using ClustalX Software F-4
1. Obtain ClustalX Software F-4
2. Create a Multi-Aligned FASTA File F-4
3. Edit Characters in the .gde File F-6
4. Create Your Library in SeqScape Software F-7

Appendix G Software Warranty

Computer Configuration G-1
Limited Product Warranty G-1
Limited Warranty G-1
Warranty Period Effective Date G-2
Warranty Claims G-2
Warranty Exceptions G-2
Warranty Limitations G-3

Glossary

Index

Preface

How to Use This Guide

Purpose of This Guide

The Applied Biosystems *ABI PRISM® SeqScape® Software Version 2.1 User Guide* provides step-by-step instructions to use this software.

Audience

This guide is intended for novice and experienced analysts and scientists who are doing resequencing.

Assumptions

This manual uses conventions and terminology that assume a working knowledge of the Windows® operating system, the Internet, and Web-based browsers.

What You Should Know Before Getting Started

To make the best use of SeqScape® Software Version 2.1 and the documentation, be sure you are familiar with:

- Microsoft® XP with Service Pack 1 or Microsoft® Windows 2000® with Service Pack 3 operating system
- The Internet and Web browser terminology
- DNA sequence detection and analysis methods
- DNA and amino acid coding conventions

Conventions Used in This Guide

Text Conventions

This guide uses the following text conventions:

- **Bold** indicates user action. For example:

  Type **0**, then press **Enter** for each of the remaining fields.
- Titles of documents and CDs are shown in italics. For example:

  ABI PRISM® SeqScape® Software Version 2.1 User Guide
- *Italic* text indicates new or important words and is also used for emphasis.
- A right arrow bracket (>) separates successive commands you select from a drop-down or shortcut menu. For example:

  Select **File > Open Project**.

  Right-click the sample row, then select **View Filter > View All Runs**.

File Naming Convention

Some alphanumeric characters are not valid for user names or file names. The characters that are illegal are listed below:

spaces

\ / : * ?“ < > |

User Attention Words

Two user attention words appear in Applied Biosystems user documentation. Each word implies a particular level of observation or action as described below:

Note: Provides information that may be of interest or help but is not critical to the use of the product.

IMPORTANT! Provides information that is necessary for proper software operation.

Examples of the user attention words appear below:

Note: Names for Reference Segments are not editable.

IMPORTANT! Do not click OK until you have completed the RDG.

Safety Conventions

Safety Alert Words

Four safety alert words appear in Applied Biosystems user documentation at points in the document where you need to be aware of relevant hazards. Each alert word implies a particular level of observation or action, as defined below:

Definitions

IMPORTANT! – Indicates information that is necessary for proper instrument operation, accurate chemistry kit use, or safe use of a chemical.

– Indicates a potentially hazardous situation that, if not avoided, may result in minor or moderate injury. It may also be used to alert against unsafe practices.

– Indicates a potentially hazardous situation that, if not avoided, could result in death or serious injury.

– Indicates an imminently hazardous situation that, if not avoided, will result in death or serious injury. This signal word is to be limited to the most extreme situations.

Examples

The following examples show the use of IMPORTANT, CAUTION, and WARNING safety alert words:

IMPORTANT! The sample name, run folder name, and path name, *combined*, can contain no more than 250 characters.

MUSCULOSKELETAL AND REPETITIVE MOTION HAZARD. These hazards are caused by potential risk factors that include but are not limited to repetitive motion, awkward posture, forceful exertion, holding static unhealthy positions, contact pressure, and other workstation environmental factors.

Do not attempt to lift or move the computer or the monitor without the assistance of others. Depending on the weight of the computer and/or the monitor, moving them may require two or more people.

Workstation Safety

Correct ergonomic configuration of your workstation can reduce or prevent effects such as fatigue, pain, and strain. Minimize or eliminate these effects by configuring your workstation to promote neutral or relaxed working positions.

CAUTION **MUSCULOSKELETAL AND REPETITIVE MOTION HAZARD**. These hazards are caused by potential risk factors that include but are not limited to repetitive motion, awkward posture, forceful exertion, holding static unhealthy positions, contact pressure, and other workstation environmental factors.

To minimize musculoskeletal and repetitive motion risks:

- Use equipment that comfortably supports you in neutral working positions and allows adequate accessibility to the keyboard, monitor, and mouse.
- Position the keyboard, mouse, and monitor to promote relaxed body and head postures.

How to Obtain More Information

Related Documentation

The following related documents are shipped with the software:

- SeqScape Online Help – Provides procedures for common tasks. Help is available from the Help menu in the main SeqScape window, or by pressing F1.
- *ABI PRISM® SeqScape® Software Tutorial*
- *ABI PRISM® SeqScape® Software Quick Reference Card*

Portable document format (PDF) versions of the documents listed above are available on the SeqScape software installation CD or by selecting **Programs** > **Applied Biosystems** > **SeqScape**. If you do not have Acrobat Reader installed on your computer, install it from the SeqScape software CD, so you can open the pdf files.

Note: For additional documentation, see "How to Obtain Services and Support" on page xix.

Send Us Your Comments

Applied Biosystems welcomes your comments and suggestions for improving its user documents. You can e-mail your comments to: **techpubs@appliedbiosystems.com**

How to Obtain Services and Support

For the latest services and support information for all locations, go to **http://www.appliedbiosystems.com**, then click the **Support** tab.

At the Support page, you can:

- Search through frequently asked questions (FAQs)
- Submit a question directly to Technical Support
- Order Applied Biosystems user documents, MSDSs, certificates of analysis, and other related documents
- Download PDF documents
- Obtain information about customer training
- Download software updates and patches

In addition, the Support page provides access to worldwide telephone and fax numbers to contact Applied Biosystems Technical Support and Sales facilities.

Introduction to ABI PRISM SeqScape Software

1

This chapter contains:

New Features in SeqScape Software v2.1 1-2
For Users of Earlier Versions of SeqScape Software 1-3
About SeqScape Software 1-5

New Features in SeqScape Software v2.1

New Features for v2.1

ABI PRISM® SeqScape® Software Version 2.1 includes the following new features and enhancements:

- Optimization
  - Faster project load, import, and save times
  - Analysis pipeline time is 33% faster
- Data analysis
  - **Automatic analysis.** Integration for automatic analysis with Applied Biosystems 3730/3730*xl* Data Collection Software version 2.0 and ABI PRISM® 3100/3100-*Avant* Data Collection Software version 2.0
  - **M13 primer trimming.** Trims the M13 universal primer sequence from the PCR product
  - **3' clear range trimming.** Trims PCR primer sequences
  - **New Basecallers.** KB basecallers identify pure and mixed bases, then assign quality values for data generated on ABI PRISM® 310, 3100, and 3100-Avant systems and Applied Biosystems 3730 and 3730xl systems
- Microsoft® Windows® XP SP1 OS compatibility
- Reports
  - **NT Mutations report.** Shows the amino acid effect of each mutation and hyperlinks link the nucleotide variant with the amino acid variant (and vice versa).
  - **New Genotyping report.** Identifies the genotype at one or more positions of all the specimens in the project and provides coverage of the samples against the expected coverage based on the targeted sequence. See page 7-39.
  - **Specimen Statistics report.** Provides the position of the clear range on the reference of the samples on the segment
  - **Amino Acid report.** Provides the option of displaying the amino acid full name or three-letter symbol
- Specimen heterozygous indel mutation (HIM) identification
  - Mutations report lists HIMs
- Amino acid alignment can be coupled to the nucleotide alignment
- Sequence Collector/BioLIMS software is no longer supported

For Users of Earlier Versions of SeqScape Software

Features of v2.0

The following features were introduced in SeqScape Software v2.0:

- **Extended Reference Data Group.** SeqScape Software v2.0 contains an extended reference data group (RDG). The RDG contains a known reference sequence and any known nucleotide or amino acid variants. The RDG available in this new version of the software enables analysis of simple or complex projects.

  The Reference Sequence within the RDG can be a:

  – Contiguous reference sequence with a single reading frame.

  – Contiguous reference sequence with multiple reading frames.

  – Reference sequence constructed from several reference segments. Each segment can come from different locations in the genome.

  The reference sequence can contain features such as exons, introns, splice junctions, primer-binding sites, and promoter regions.
- **Heterozygous insertion/deletion mutation.** SeqScape identifies potential instances of this variant which often require manual review by trained personnel.
- **Library searching.** You can compare each consensus sequence to a sequence library to identify the closest match genotype, allele or haplotype.
- **Enhanced reports.** You can customize reports. Each variant in the report is hyperlinked to the sequence data, providing rapid transition from results to data. The results reports eliminate the need to manually record results. You can automatically sort and reorganize any report.
- **Password protection and audit trail.** The software protects your data by providing password protection, automatic lockout when the software is inactive, and three levels of access control. An audit trail records each manual insertion, deletion, or base modification, with reasons for each change.
- **Integration automation.** The software uses an improved process for setting up samples for automated analysis on Applied Biosystems 3730/3730*xl* instruments.

- **New Basecallers.** The KB basecaller is a new algorithm that identifies mixed or pure bases and generates sample quality values. The ABI basecaller is an algorithm used in sequencing analysis software.
- **Basecalling with ABI basecaller only is no longer available.** In SeqScape Software v1.1, you can choose to basecall data with ABI basecaller or ABI basecaller with TraceTuner™ Software. In software v2.0, you do not have the option to basecall with ABI basecaller only. The new options are:
  - Basecall with ABI and TraceTuner software (automatic)
  - Basecall with KB basecaller
- **Implicit Reference is no longer available.** In software v1.1, you can have an empty RDG and use the first specimen as your implicit reference sequence, but this is no longer available in v2.0. However, you can create an RDG and add an .ab1 sample file as a reference sequence.

About SeqScape Software

Genetic Analyzer Applications

SeqScape software is one of a suite of Applied Biosystems Genetic Analyzer software applications designed to control an instrument, collect data, and manage automated analysis. This suite of data collection and analysis software systems includes:

- GeneMapper™ Software – Performs genotyping using fragment analysis methods.
- ABI PRISM® Sequencing Analysis Software – Displays, analyzes, edits, and prints sequencing files.
- ABI PRISM® SeqScape® Software – Performs sequence comparisons for variants identification, SNP discovery, and SNP validation.

SeqScape Software Applications

Common resequencing applications include:

- SNP discovery and validation
- Mutation analysis and heterozygote identification
- Sequence confirmation for mutagenesis or clone-construct confirmation studies
- Identification of genotype, allele, and haplotype from a library of known sequences

Resequencing Data with SeqScape Software

SeqScape software allows analysis of resequenced data, comparing consensus sequences to a known reference sequence and optionally searching against a sequence library.

For example, a simple project might contain one contiguous reference sequence in a single reading frame, with no known nucleotide or amino acid variant information. SeqScape software compares a consensus sequence to this reference sequence, identifying any differences.

A more complex project might include a reference sequence constructed from several reference segments representing multiple exons and introns. You can use SeqScape software to:

- Build unique sequence layers composed of different groupings of reference sequence features.
- Compare consensus sequences to each unique layer.
- Identify differences.
- Compare the sequence to a library of sequences to identify the closest match.

Data Sources for Resequencing Projects

You can create projects in SeqScape software using sequencing data generated from the following systems:

- ABI PRISM® 310 Genetic Analyzer
- ABI PRISM® 377 DNA Sequencer
- ABI PRISM® 3100-*Avant* Genetic Analyzer
- ABI PRISM® 3100 Genetic Analyzer
- ABI PRISM® 3700 DNA Analyzer
- Applied Biosystems 3730 DNA Analyzer
- Applied Biosystems 3730*xl* DNA Analyzer

Each project can contain:

- Unanalyzed sample files (.ab1)
- Previously basecalled sample files (.ab1)
- Text sequences (.seq or .fsta)
- Aligned consensus sequences

A single project can contain sample files from one or a mixture of instrument platforms. The software analyzes the data, displays several views of the analyzed project, and reports on results for quality control and data review.

Levels of Automated Analysis

SeqScape software performs two levels of analysis:

- It identifies variants, positions that differ from the reference sequence, and classifies those variants as known or unknown.
- It searches a library of alleles or haplotypes to identify the alleles that most closely match the sample.

What the Software Does

When you have added a reference sequence, a library, and sample files, SeqScape software performs two levels of analysis:

- Identification of nucleotide and amino acid variants.The software identifies positions that differ from the reference sequence and classifies those variants as known or unknown variants.
- Identification of genotypes, alleles, or haplotypes from a library. In addition to identification of variants, the software searches a library of genotypes, alleles, or haplotypes and identifies the alleles that most closely match each consensus sequence.

How the Software Performs Analyses

You provide the following information to the system before analysis:

- A reference sequence (backbone) made up of one or more reference segments and any known nucleotide variant information or amino acid variant information. (SeqScape software uses the backbone to classify all polymorphic positions as known variants or unknown variants.)
- An allele library (a set of sequences for the alleles or haplotypes).

Using the reference sequence, variants, allele library, and software settings, you create a reusable project template. With this template and the sequencing samples, SeqScape software:

- Performs basecalling, quality value assignment, and mixed base identification, in that order.
- Trims low-quality bases from each sequence.
- Identifies poor-quality samples and removes them from further analysis.
- Assembles the remaining samples against the reference sequence and generates a specimen consensus sequence.
- Reviews the basecalling quality values and the sample assembly to confirm, improve, and assign quality values to the consensus sequence.
- Identifies variants by aligning specimen sequences to the reference sequence and comparing the specimen consensus sequences to the reference sequence.
- Generates nine detailed reports.

Note: If you link a library to a project, the software also automatically searches the library to find the closest match to each consensus sequence.

When the analysis is complete, the software generates a project file that contains sample files, a consensus sequence for each specimen, and nine reports. You can print and export your results.

Getting Started 2

This chapter contains:

Administrator: Registering the Software 2-2
Hardware and Software Requirements 2-3
SeqScape Software Versions 2-5
Preparing to Install SeqScape Software 2-6
Installing the Full Version for the First Time 2-7
Upgrading from SeqScape Software Versions 1.0, 1.1, or 2.0 2-8
Removing SeqScape Software 2-10
Starting SeqScape Software for the First Time 2-11
New Users Logging In for the First Time 2-22
SeqScape Software Structure 2-23
SeqScape Software Toolbar 2-25
Menus on the Main SeqScape Window 2-27
Workflow 2-29

Administrator: Registering the Software

This chapter provides information you need to know before installing and using the ABI PRISM® SeqScape® Software Version 2.1. The administrator must follow the procedures in this section, "Administrator: Registering the Software," through "Starting SeqScape Software for the First Time" on page 2-10.

License and Warranty

Before you begin, read Appendix F, "Software Warranty Information." Appendix F explains your rights and responsibilities regarding the SeqScape software.

During the installation process of the software, you must accept the terms and conditions of the Software License Agreement before the software can be installed.

Registering Your Software

To register your copy of SeqScape software, complete the registration card (included in this software package) and return it to Applied Biosystems.

Registering the software enables Applied Biosystems to send you notification of software updates and any other future information that may be specific to SeqScape software owners.

IMPORTANT! Your product registration number is located on the registration card. Be sure to record the number here before you return the registration card.

Registration Number:

Hardware and Software Requirements

Minimum System Requirements

SeqScape software can be installed on a computer that meets the minimum requirements summarized in Table 2-1 below.

In general, the more memory and the more processing power in the system, the better its performance.

Table 2-1 Minimum System Requirements

System Component	Minimum Requirements
CPU	733 MHz or faster with an Intel Pentium® III or IV processor The software does not run on computers with a dual processor or with a Xeon chip set.
CD-ROM drive	Any
Operating system	Microsoft® Windows 2000® OS with Service Pack 3 Microsoft Windows® XP with Service Pack 1
RAM	256 MB. 512 MB is recommended
Printer	An HP® 4500, 8100, 990cxi or an Epson® 980 printer is recommended
Monitor	A 17-inch monitor or larger is recommended A monitor with 1024 × 768 resolution is recommended
Disk space	1 GB Storage requirements depend primarily on the quantity of data to be generated and stored. It is common to store many SeqScape software project files on the analysis computer. Install SeqScape software on a partition with enough space for the projects and their files because SeqScape software stores data files in the area where the program is installed.

Hard Drive Partitions

The installer uses the following location for SeqScape software files:

drive letter:\Applied Biosystems\SeqScape

The drive letter is determined by the following conditions:

Table 2-2 Drive Letter Determination:

If the computer …	The installer selects drive …
is not connected to a genetic analyzer	D (default), or C (if D drive is not available)
has data collection software that is connected to a genetic analyzer	E

SeqScape Software Versions

There are four versions of SeqScape Software v2.1 available:

Software Version	Intended for a system that ...	Existing Data
SeqScape Software v2.1 full release version	Has no SeqScape software installed, or has the 30-day demo version installed	Data created in v1.0, 1.1, or 2.0 will be moved to version 2.1.
SeqScape Software v2.1 upgrade	Has SeqScape software v1.0, 1.1, or 2.0 installed	Data created in v1.0, 1.1, or 2.0 will be moved to version 2.1.
SeqScape Software v2.1 update	Has SeqScape v2.0 installed	Data created or saved in v2.0 will be moved to v2.1. **IMPORTANT!** This update does not work on systems with SeqScape software v1.0 or 1.1.
SeqScape Software v2.1 30-day demo	Meets the minimum requirements	You must export any files created in the 30-day demo version before the software expires. You can then import any exported files into the full release version of SeqScape v2.1.

Preparing to Install SeqScape Software

Note: An administrator should install the software and use it for the first time. The administrator can set up the software for the analyst, scientist, or other administrator users.

To prepare for the installation:

1. Ensure that your system meets the minimum requirements (see "Hardware and Software Requirements" on page 2-3).

   Check that you have at least 1 GB of free disk space to accommodate SeqScape software, and sufficient space for all projects and their sample files.

2. Temporarily turn off any virus protection software.

3. Exit all programs except the Applied Biosystems 3730/3730*xl* or ABI PRISM® 3100/3100-*Avant* Data Collection software, if applicable.

IMPORTANT! To properly install SeqScape Software v2.1 on a computer that is connected to Applied Biosystems 3730/3730*xl* DNA Analyzers or ABI PRISM® 3100/3100-*Avant* Genetic Analyzers, the **data collection software must be running**. SeqScape software will not register with the Data Service if the data collection software is not running. See Chapter 11, "Automating Analysis," for more information on file sharing and automation.

Installing the Full Version for the First Time

The administrator of the software installs the software and sets up new users.

Follow this procedure to install the full version of SeqScape Software v2.1 only if you are installing on:

- A computer that does not have a previous version of the software, or
- A computer that has a 30-day demo version of the software installed

IMPORTANT! If you are using a SeqScape software 30-day demo, export your data **before** the end of the 30 days and **before** you load the full version of SeqScape software. To ensure data is not lost:

a. Export your data: **Select Tools** > **SeqScape Manager** > **Target Files** and click **Export**.

b. Load the software as described in the procedure below.

c. Re-import your data.

To install the full version of SeqScape software for the first time:

1. Insert the *ABI PRISM® SeqScape® Software v2.1* CD into the computer CD-ROM drive.
2. If the installer does not start automatically, double-click **setup.exe** on the CD.
3. When the InstallShield Wizard Complete window opens, click **Finish**.

After the software is installed, the administrator must log into the software for the first time. After the initial login, the software can be set up for additional users.

Upgrading from SeqScape Software Versions 1.0, 1.1, or 2.0

About the Upgrade

The SeqScape Software v2.1 installer automatically:

- Detects the previous version and backs up your data folder
- Removes the previous version, then installs the new version of SeqScape software

Follow the instructions on the installer.

Upgrading to v2.1

To upgrade your software:

1. Insert the appropriate SeqScape software installation CD into the CD-ROM drive:

If you are upgrading from SeqScape software version	Then, insert the CD titled
2.0	*ABI PRISM® SeqScape® Software v2.1* ***Update*** CD
1.0 or 1.1	*ABI PRISM® SeqScape® Software v2.1* ***Upgrade*** CD

2. If the installer does not start automatically, double-click **setup.exe**.
3. In the SeqScape Registration Code dialog box, enter your registration code for SeqScape software v2.0.
4. Follow the instructions to upgrade the software.

After the software is installed, you must register and log in as an Administrator user. After you log in as Administrator, you can set up additional users with Admin, Scientist, or Analyst permissions.

Existing Users

All existing users of an earlier version of SeqScape software will have Analyst privileges. Only a user belonging to the Administrator group can change the user to Scientist or Analyst. A dialog box opens for users who existed in previous versions to set up their user profiles (name and password) when they try to use SeqScape Software v2.1 for the first time.

Removing SeqScape Software

What the Uninstallation Process Does

To completely remove SeqScape software from your computer, follow the procedure in this section. The uninstallation process:

- Deletes all folders and files installed by SeqScape software. However, if you moved SeqScape software folders or files from their original installed location, they may not be found and deleted by the uninstallation process.
- Does not delete any files or folders created by users. Any files that have been added to the application folders, such as those created when the applications are run, are not deleted by the uninstallation process.

Removing the Software

To uninstall SeqScape software:

1. Select **Start** > **Programs** > **Applied Biosystems** > **SeqScape** > **Uninstall SeqScape v*x.x***.
2. Continue to follow the instructions to uninstall the software.

   When the uninstallation is complete, all the software program files are removed. Your data files remain on the computer. The uninstaller does not delete any folders or files created after installation. If you want to delete any folders and files created after installation, you must remove them manually.

Starting SeqScape Software for the First Time

Before You Begin

When you start the software for the first time, you are prompted with a registration dialog box that creates an administrator account. Log in to SeqScape software as `Admin` and enter the password you created.

To create new users, you must log in as `Admin`. Logging in with a user name allows SeqScape software to track each user's interactions with each project.

For information on the privileges for each category of user using the software, refer to Appendix D, "User Privileges."

File-Naming Convention

Some alphanumeric characters are not valid for user names or file names. The invalid characters are below:

spaces

\ / : * ? " < > |

An error message is displayed if you use any of these characters. You must remove the invalid character to continue.

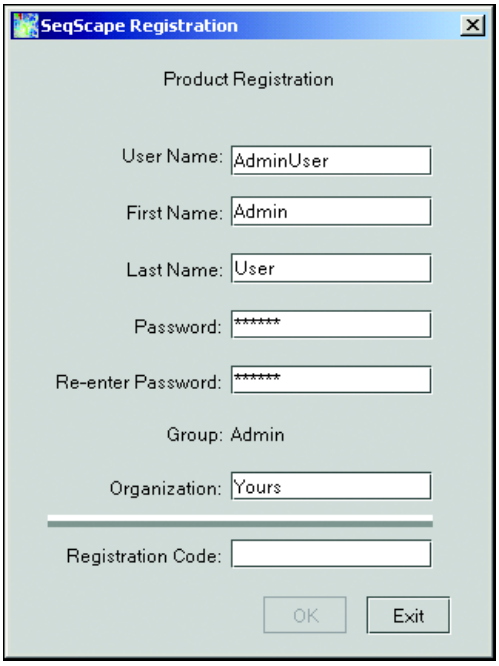
IMPORTANT! User names cannot be named seqscape_admin in this version of the software. If you have used this user name in a previous version of the software, you must change the user name to follow the File Naming Convention shown above.

Starting SeqScape Software

To start the software for the first time:

1. Double-click the SeqScape software desktop shortcut.

   The SeqScape Registration dialog box opens.



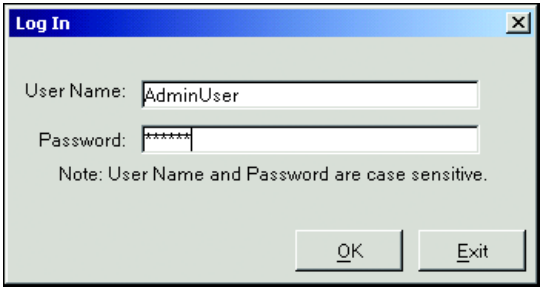
2. In the **SeqScape Registration** dialog box, enter all the information in the text fields. The User Name and password must be 6 to 15 characters long.

   The first user created is automatically assigned Administrator privileges.

3. Enter the registration code on the registration card you received with your software.

4. Click **OK**.

   The splash screen appears while the program is loading, then the Log In dialog box opens.



5. Enter your user name and password again, then click **OK**.

   The License dialog box opens.

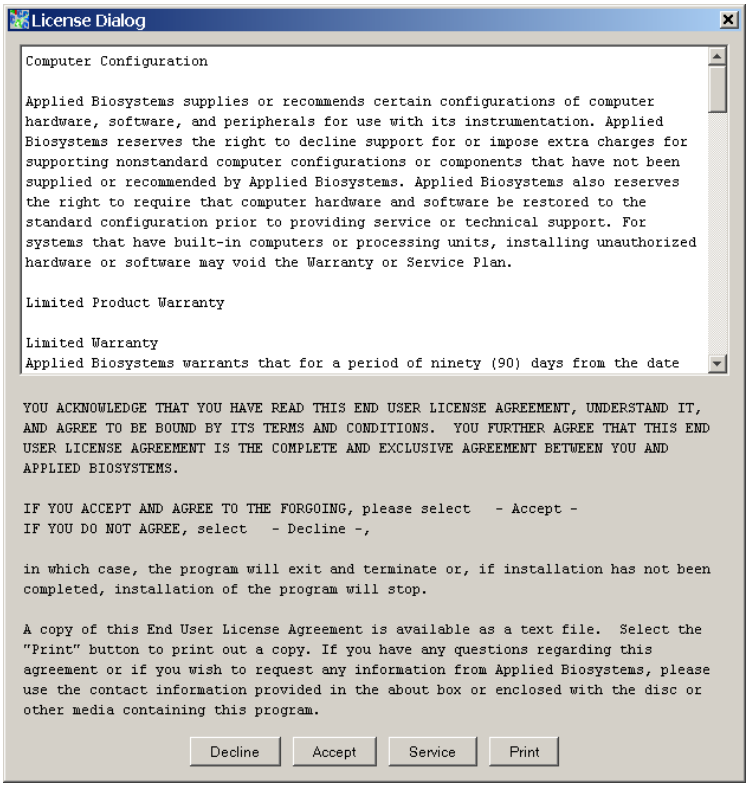
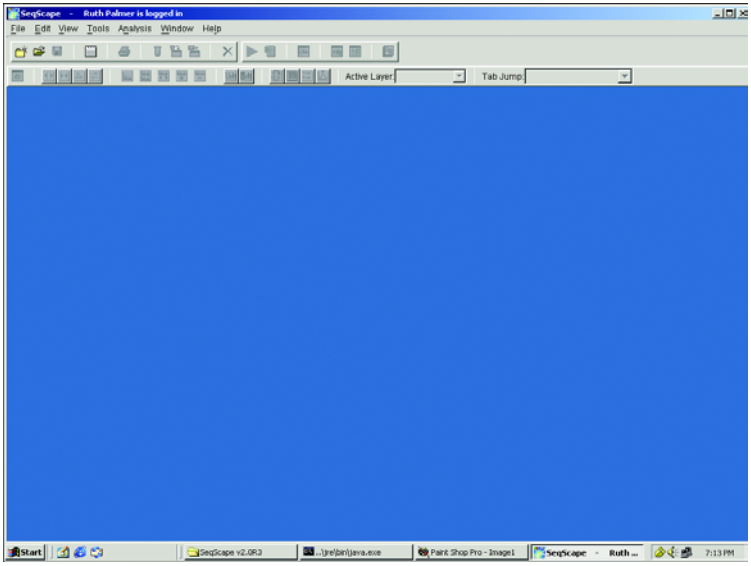

License Dialog
Computer Configuration
Applied Biosystems supplies or recommends certain configurations of computer hardware, software, and peripherals for use with its instrumentation. Applied Biosystems reserves the right to decline support for or impose extra charges for supporting nonstandard computer configurations or components that have not been supplied or recommended by Applied Biosystems. Applied Biosystems also reserves the right to require that computer hardware and software be restored to the standard configuration prior to providing service or technical support. For systems that have built-in computers or processing units, installing unauthorized hardware or software may void the Warranty or Service Plan.
Limited Product Warranty
Limited Warranty
Applied Biosystems warrants that for a period of ninety (90) days from the date
YOU ACKNOWLEDGE THAT YOU HAVE READ THIS END USER LICENSE AGREEMENT, UNDERSTAND IT, AND AGREE TO BE BOUND BY ITS TERMS AND CONDITIONS. YOU FURTHER AGREE THAT THIS END USER LICENSE AGREEMENT IS THE COMPLETE AND EXCLUSIVE AGREEMENT BETWEEN YOU AND APPLIED BIOSYSTEMS.
IF YOU ACCEPT AND AGREE TO THE FORGOING, please select - Accept -
IF YOU DO NOT AGREE, select - Decline -,
in which case, the program will exit and terminate or, if installation has not been completed, installation of the program will stop.
A copy of this End User License Agreement is available as a text file. Select the "Print" button to print out a copy. If you have any questions regarding this agreement or if you wish to request any information from Applied Biosystems, please use the contact information provided in the about box or enclosed with the disc or other media containing this program.
Decline
Accept
Service
Print

Figure 2-1 License Dialog

6. Read the license agreement then click **Accept**.

The SeqScape software main window opens.



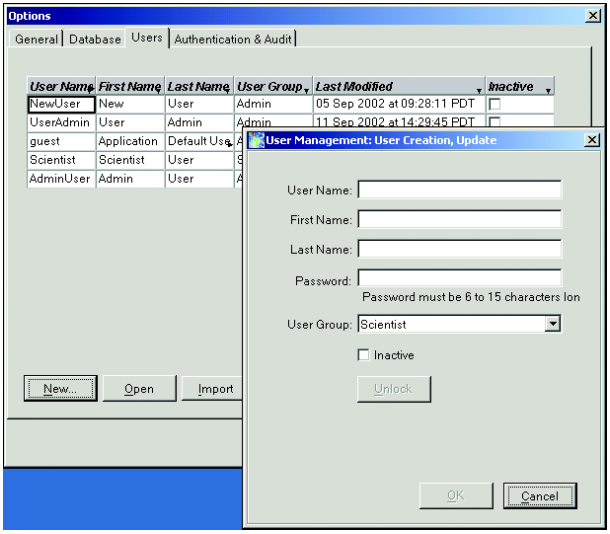
Creating New Users

Because SeqScape software tracks the projects and settings for each user, Applied Biosystems recommends that you create users for each individual who uses SeqScape software on the computer. The Users tab allows exporting of user names and access privileges for these users.

IMPORTANT! The administrator is the only person who can set up and change the information in the Users tab. The selections in this tab are inactive for all other users.

To set up new users:

1. Select **Tools** > **Options** to open the Options dialog box.
2. In the Options dialog box, select the **Users** tab, then click **New**.



3. Fill in the appropriate user name, password, first and last names, then select the level of user from the Group drop-down list.

   Note: Enter a User Name that contains only alphanumeric characters. This field must not contain any spaces or characters that do not conform with the Microsoft® Windows OS file system. Refer to “File-Naming Convention” on page 2-10.

   The new user appears in the list in the Users tab.

4. To set up users on multiple computers, we recommend that the administrator do the following:
   a. Create the list of users.
   b. Export the file.
   c. Install SeqScape software on the other computers.
   d. Import the user file.

New users can log in after exiting and then restarting SeqScape software.

Setting Up Authentication & Audit

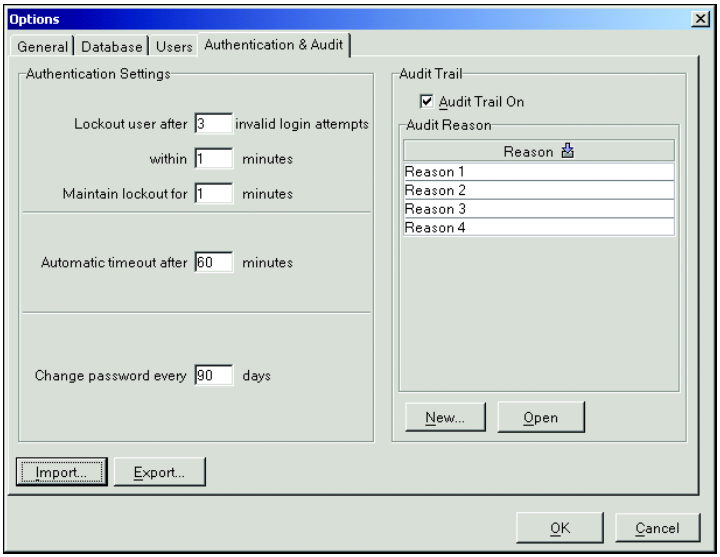
Users belonging to the Administrator group can change the default settings in the Authentication & Audit tab for security features of the application.

Note: The Administrator is the only person who can set up and change the information in the Authentication & Audit tab. The selections in this tab are inactive for all other users.

The Authentication & Audit panes provide a way to track the changes in projects such as base change, variants, or processes you want to track. You must turn Audit Trail On for tracking to occur.

To set up authentication and auditing:

1. Select the **Authentication & Audit** tab to change the defaults for the Authentication Settings:



   a. Lockout occurs when a user enters an incorrect password or user name the number of times you select for the **Lockout user after invalid login attempts** field. Enter the number or accept the default.

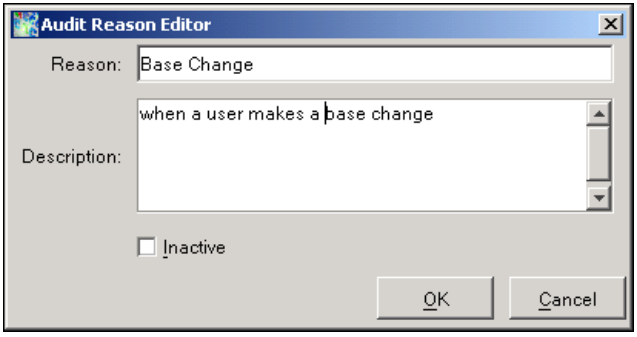
   b. The **within minutes** field indicates that the user will be locked out if the maximum number of attempts occur within the time entered in this field. Enter a number or accept the default.

c. The **Maintain lockout for minutes** field indicates the number of minutes that must elapse before the user can login again after being locked out of SeqScape software. Enter the number of minutes or accept the defaults.

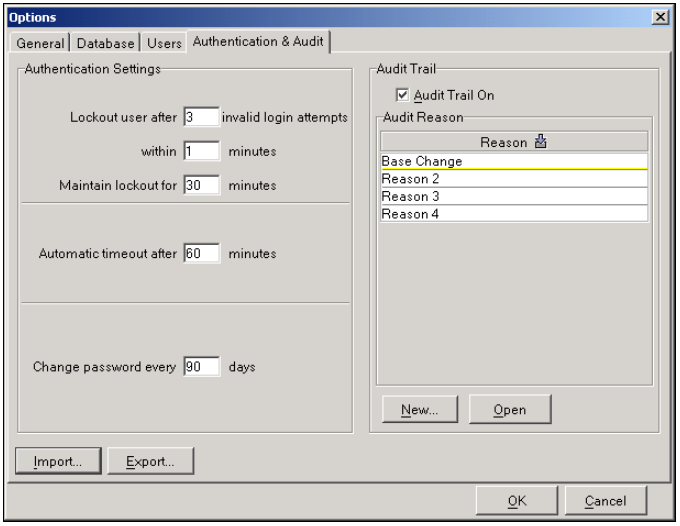
d. The **Change password every days** field indicates the number of days before the users must enter a new password. Enter a number of days or accept the default.

2. In the Audit Trail pane, select the **Audit Trail On** check box to have a dialog box open whenever an indicated reason occurs.

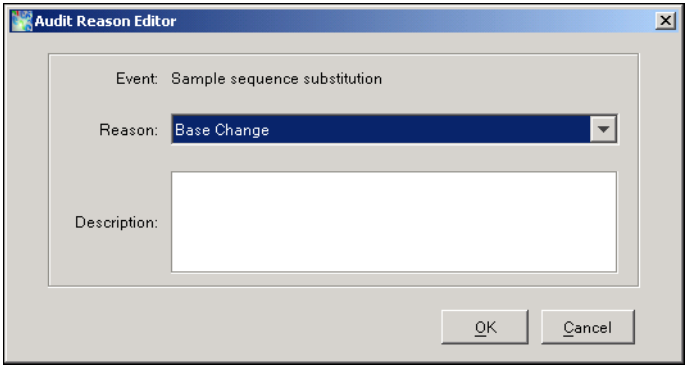
3. In the Audit Reason pane, enter reasons to provide an audit trail.
   a. Double-click the **Reason 1** field, or highlight it and click **New**.



   b. In the Reason field, type a reason for a change to the project to identify, for example, a base change, or a variant that is imported.
   c. Enter a description of the reason, if desired.
   d. Click **OK** in the Audit Reason Editor. The first reason appears in the list in the Options dialog box.



e. Whenever a change is made in any of the project views, the Audit Reason Editor dialog box opens as shown below. Select the reason for the change from the drop-down list.



4. If desired, click **Export** in the Options General tab and navigate to export the configuration settings to another computer. The Import button allows configuration settings to be imported from another computer.

5. Click **OK** in the Options dialog box to save the authentication and audit settings.

Note: It is possible to import or export Authentication & Audit configurations from one computer to another. For example, an administrator may want to set up authentication and audit information for many users, then select all the files and export them to other systems using SeqScape software.

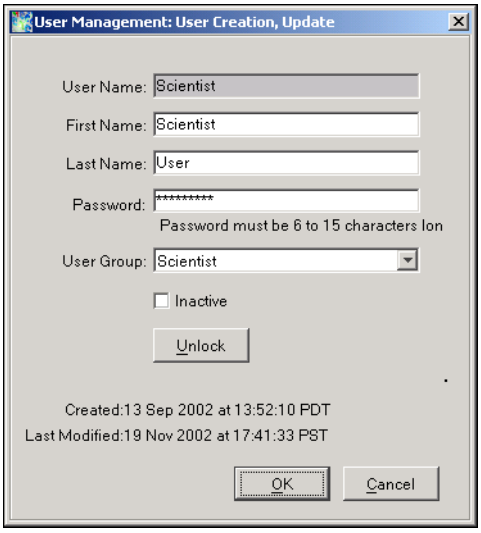
Changing User Information

You can change the default settings for a user you are setting up.

IMPORTANT! The Administrator is the only person who can set up and change the information in the Users tab. The selections in this tab are inactive for all other users.

To change any of the information for a user:

1. In the Options dialog box, select the **Users** tab.
2. Double-click the name in the list to open the User Management dialog box.



3. Change or correct the user information and click **OK**.
4. If desired, click the **Export** button in the Options dialog box to export the application configuration settings and/or settings for a single user or multiple users in a zipped .ctf format.
5. Enter the path for exporting files in the Export User dialog box, then click **Export**.
6. Click **OK** to close the Options dialog box.

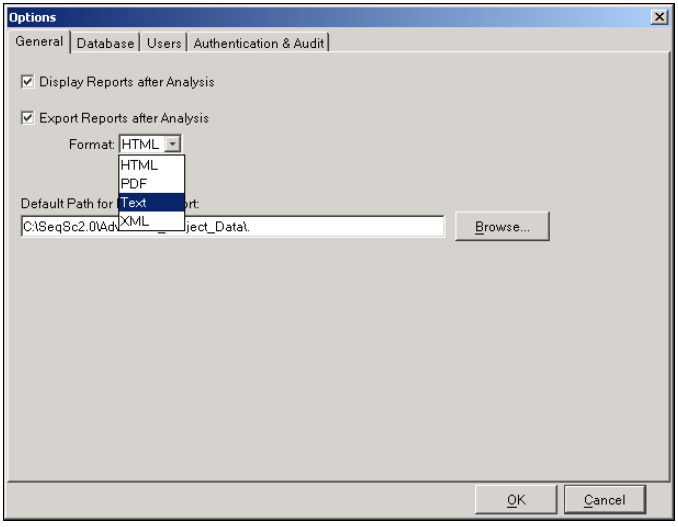
Note: This process can be used by the first administrator to set up additional users or another administrator. It is possible to import or export user settings from one computer to another. For example, an administrator can set up user information for many users, then select all the user files and export them to other systems using SeqScape software.

Setting Up the Default Directory

The default directory should be set up for users for importing and exporting data files. If the directory path is not set up, the default directory opens to C:\.

To set up the default directory path:

1. In the SeqScape software main window, select **Tools** > **Options**.
2. In the General tab, select the appropriate check boxes for your setup, if desired.
   a. Select the **Display Reports after Analysis** check box.
   b. Select the **Export Reports after Analysis** check box, if desired, then select the format in which to export them from the Format drop-down list.



3. Click **Browse** and navigate to the directory to use as the default for files to be stored.
4. Click **Open**.

   The exported reports are stored in the directory you select as the default.
5. Click **OK** to save the directory path and close the dialog box.

New Users Logging In for the First Time

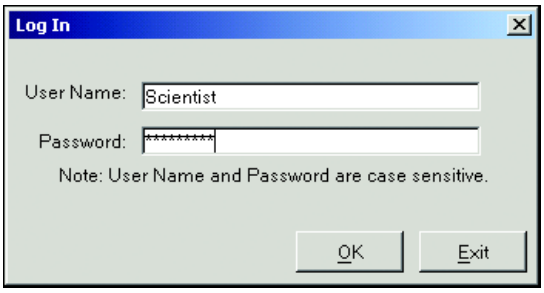
When New Users Log In

After the installation and setup are complete, new users can log in to the software.

To log in to SeqScape software:

1. Start the software by double-clicking the desktop shortcut .
2. The Log In dialog box opens, showing the last user's name. Enter your user name and password, then click **OK**.

   IMPORTANT! If you have forgotten your user name or password, contact your administrator.



SeqScape software is ready to use.

Note: All existing users of an earlier version of SeqScape software will have Analyst privileges. Only a user belonging to the Administrator group can change the user to Scientist or Analyst. Users who existed in previous versions will be asked to set up their user profiles (name and password) when they try to use SeqScape Software v2.1 for the first time.

SeqScape Software Structure

SeqScape software is organized around two main windows:

- SeqScape Manager window, from which you enter and manage the information necessary to perform analyses
- Project window, from which you manage the results of analyses

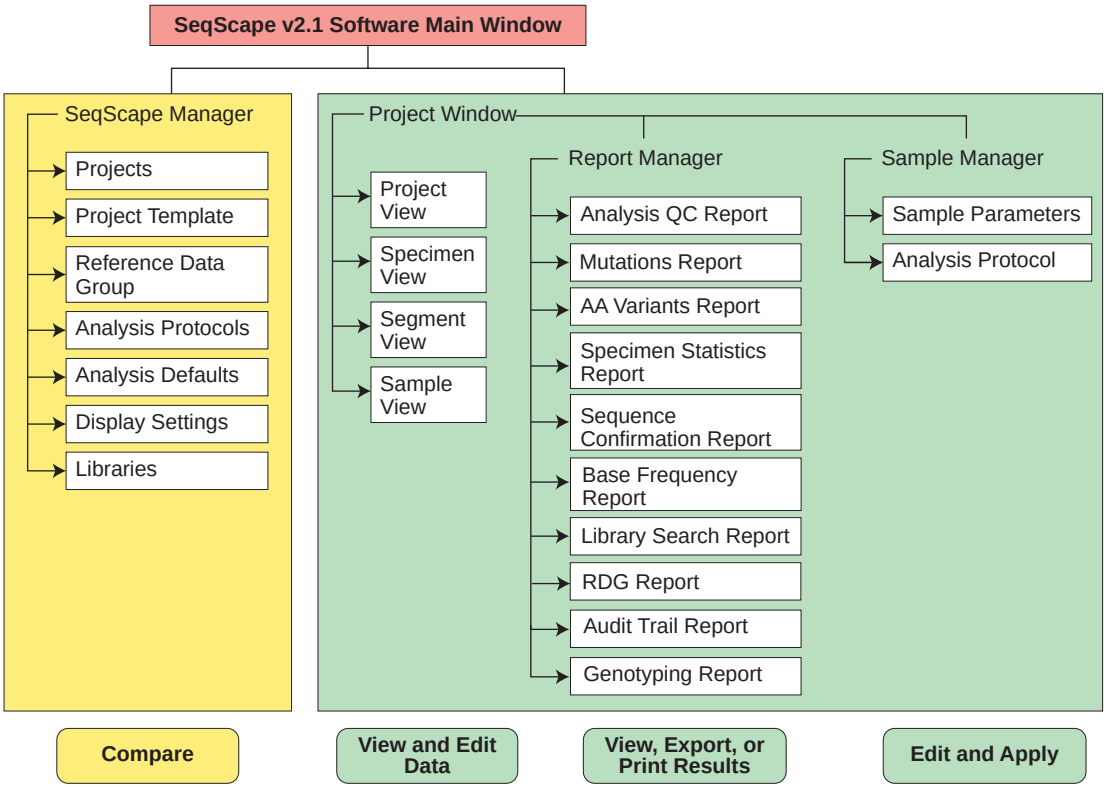


Figure 2-2 SeqScape Software Structure

SeqScape Manager Window

In the SeqScape Manager, you configure projects by creating project templates. The project templates can be reused in multiple projects and can be exported to be shared with other researchers. The project templates contain:

- Reference sequence information
- Analysis settings (including analysis protocols)
- Display settings

Project Window

In the Project window, you can view your data in the following ways:

Table 2-3 Views in the Project Window

View	Description
Project View	Shows the reference sequence, each specimen consensus sequence, and electropherogram snippets for each sample file in each specimen. The Expanded Nucleotide View shows all the nucleotides. The Collapsed Nucleotide View shows only variants of the nucleotides. The Expanded Amino Acid View shows all the amino acids. Characters (NT or AA) that are the same as the reference are shown as dots. The Character/Dots button switches to show or hide the view. The Identification pane, which shows the library search results, appears at the bottom of the Project view.
Specimen View	Shows the clear range and orientation of each sample and how they line up to the reference sequence, and the overview pane with active ROIs.
Segment View	A table of sample information. Clicking a row in the table shows the corresponding sample sequence below. The Layout tab shows the direction of each sample within the segment. The Assembly tab shows samples aligned to the consensus sequence. An overview pane that represents forward and reverse sequences, variants, and ROIs. Electropherograms can be displayed for one or all sequences.
Sample View	Shows pertinent information for the sample, which includes annotation, sequence, electropherogram and raw data.

Refer to Chapter 10, “Sample and Consensus Quality Values,” for detailed descriptions of these views.

SeqScape Software Toolbar

The SeqScape software toolbar displays buttons for software functions that you are likely to use often. Refer to the next two figures for the names, descriptions, and keyboard shortcuts for each button. The top row of buttons, Figure 2-3, are processing tools.

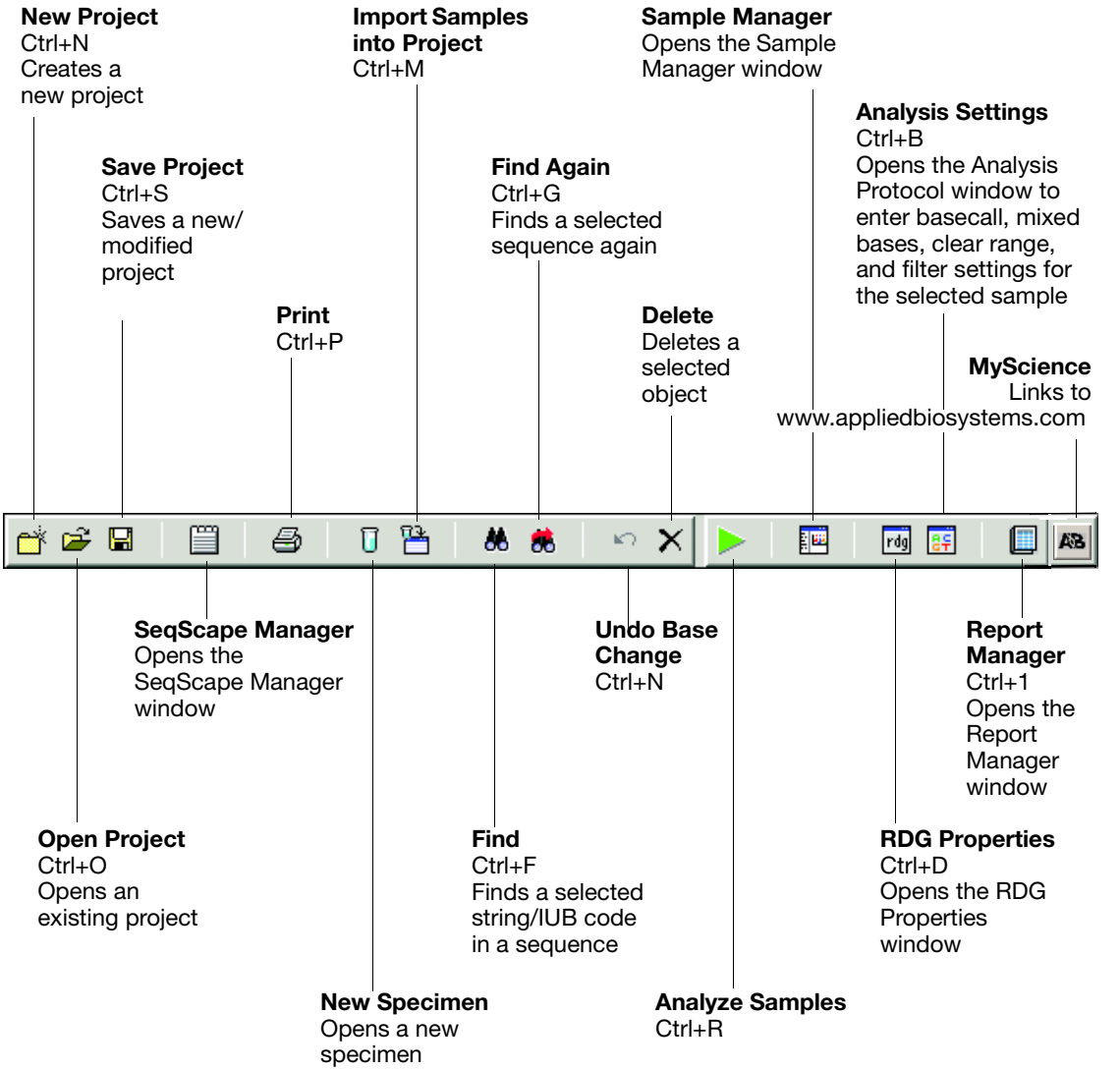


Figure 2-3 Main Toolbar

The second row of buttons, Figure 2-4, are viewing options for the projects you create.

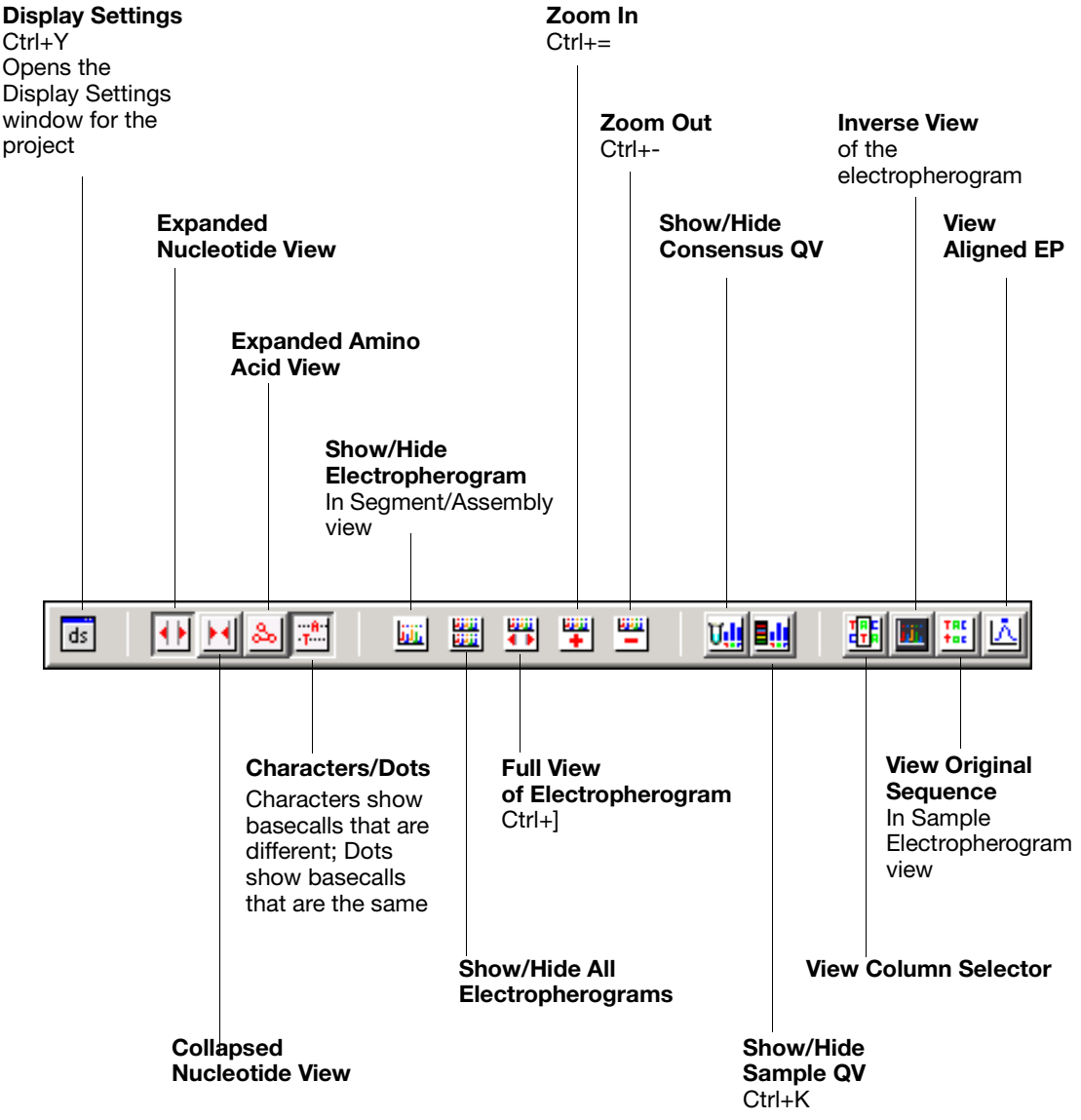
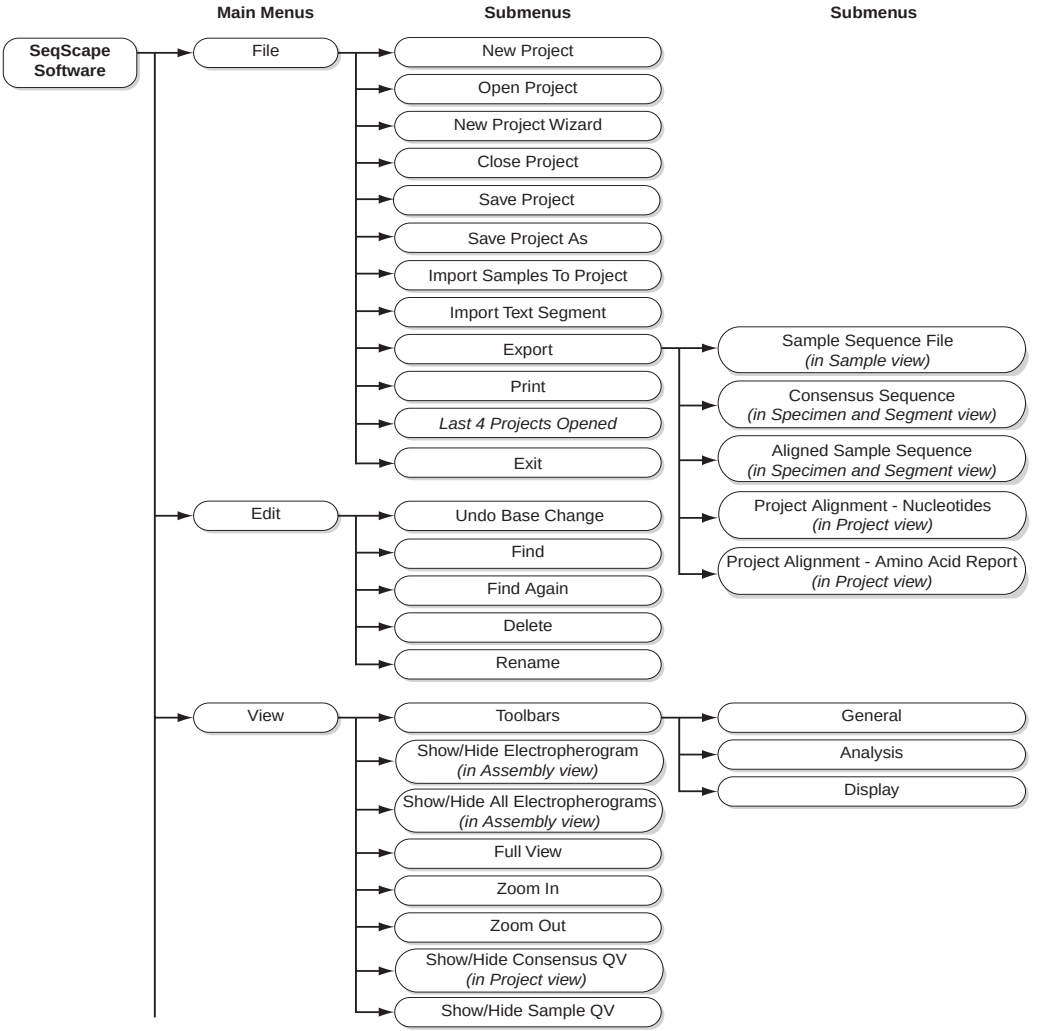


Figure 2-4 Viewing Toolbar

Menus on the Main SeqScape Window

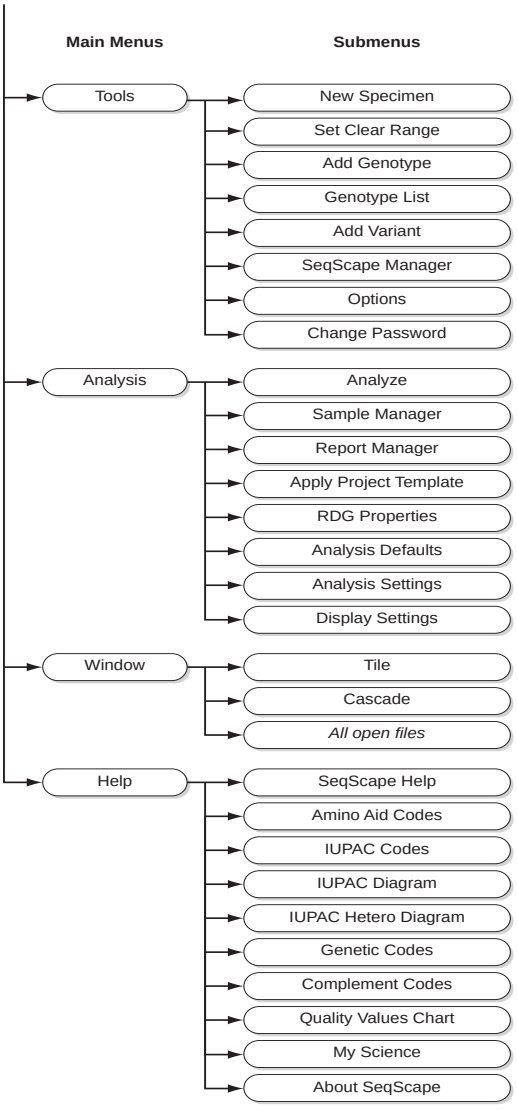
Figure 2-5 shows the menu structure of the main SeqScape window.



Continued on next page...

Figure 2-5 Main SeqScape Window Menus

Continued from previous page



Workflow

A typical workflow using SeqScape software is shown below.

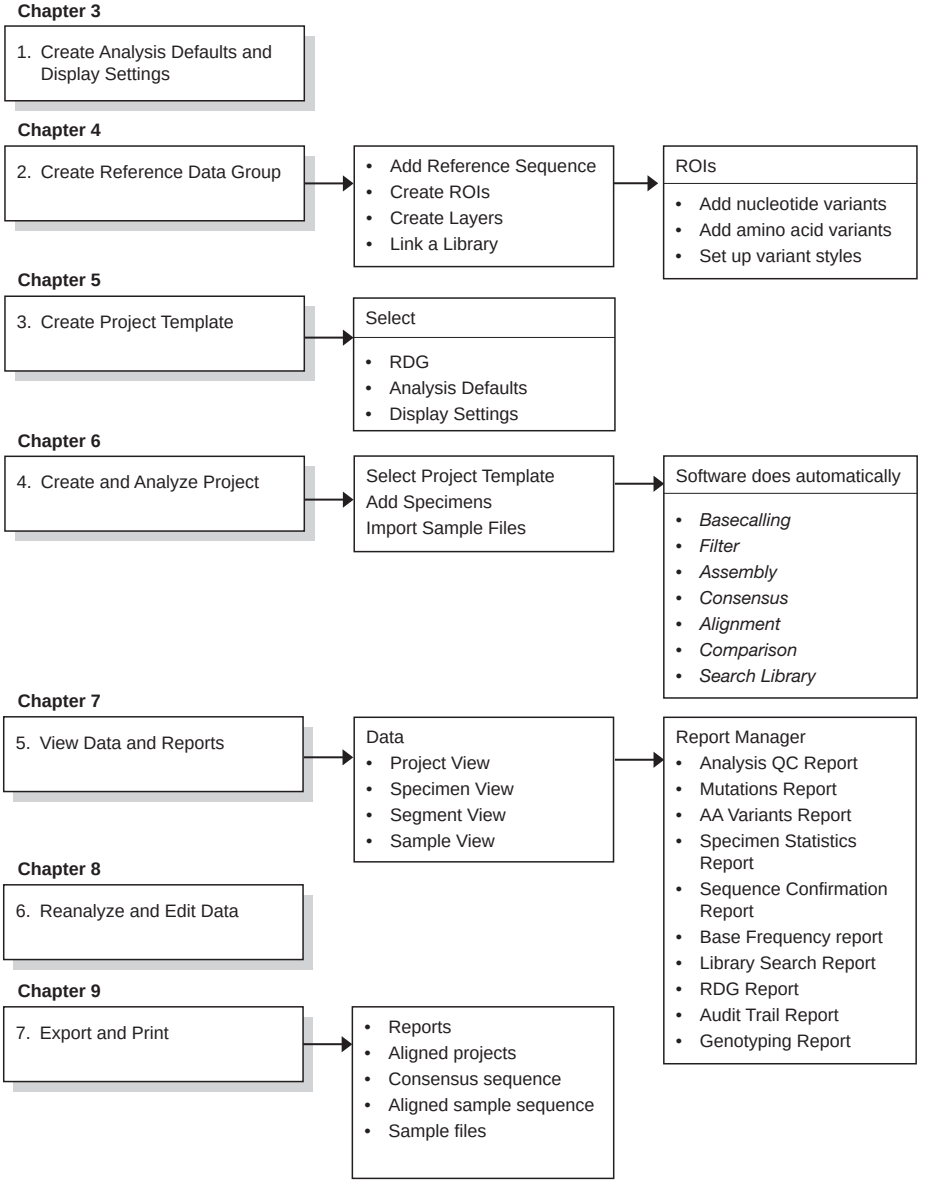


Figure 2-6 Typical Workflow for a Project

Creating Analysis Defaults and Display Settings

3

This chapter contains:

Workflow for This Chapter 3-2
Analysis Defaults Settings 3-3
Creating Analysis Protocols 3-3
Specifying the Analysis Settings. 3-12
Specifying Display Settings 3-18

Workflow for This Chapter

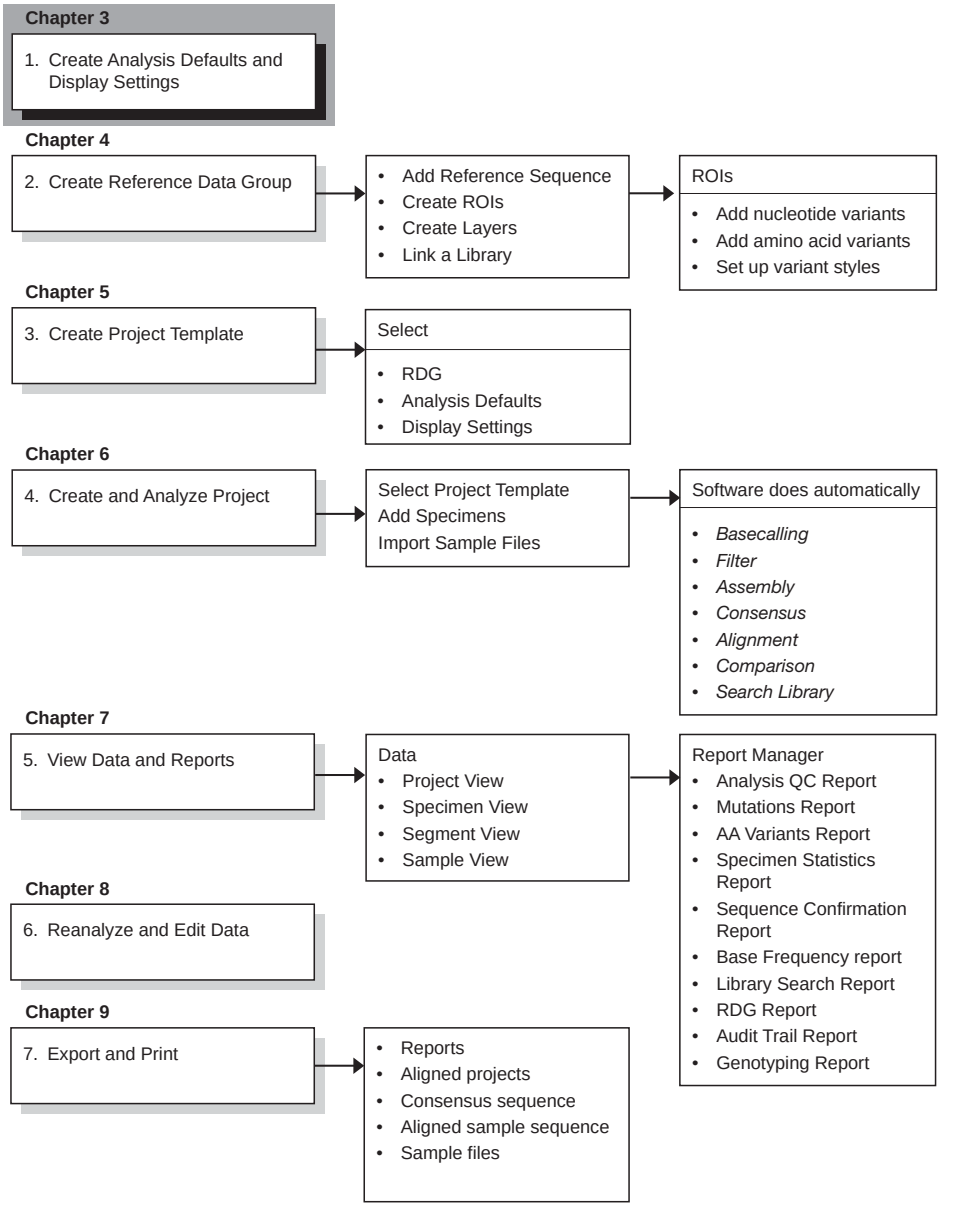


Figure 3-1 Create Analysis Defaults and Display Settings Step

Analysis Defaults Settings

Analysis Defaults are a component of the project template. They are used to set the analysis settings for all the samples as they are imported into a project.

Creating Analysis Protocols

Before you create analysis defaults, you need to create the analysis protocol. An analysis protocol in the ABI PRISM® SeqScape® Software Version 2.1 specifies the analysis conditions to be applied to your samples. You can specify the analysis protocol settings for one or more samples. You must select an analysis protocol before selecting analysis defaults. The analysis protocol settings include:

- Basecalling
- Mixed bases
- Clear range
- Filtering

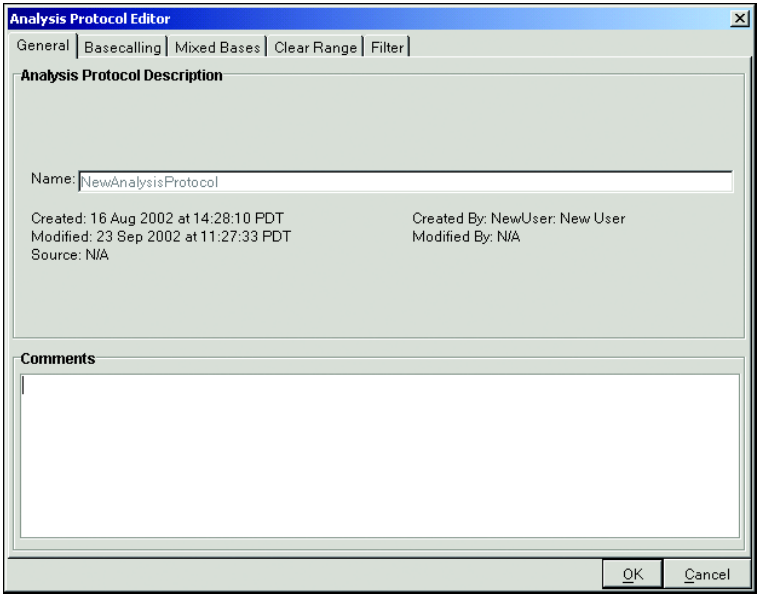


Figure 3-2 Analysis Protocol Editor Showing General Tab

Analysis Protocol Editor Tabs

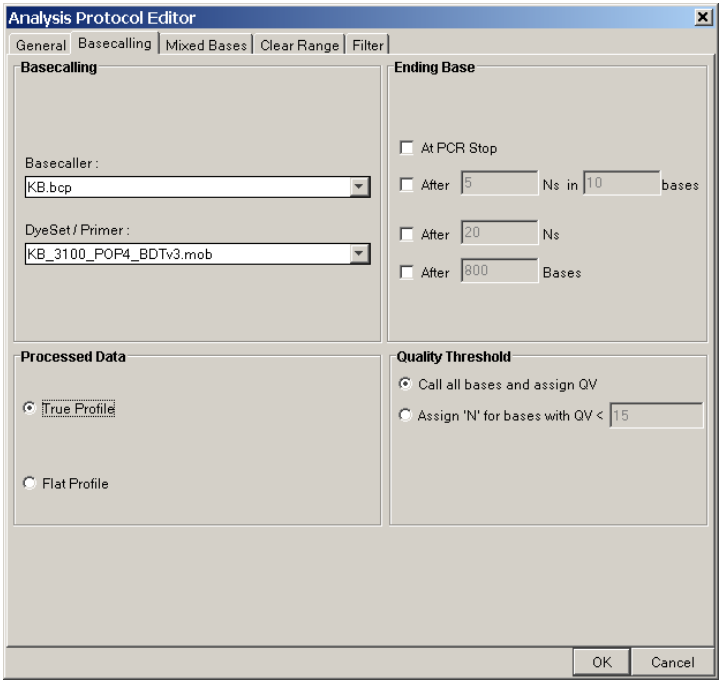
The Analysis Protocol Editor tabs and descriptions are used to set up the analysis protocol to select with the analysis defaults.

- The **General** tab contains general information on the analysis protocol, for example, the name, creation date, and modification date. Refer to Figure 3-2.
- The **Basecalling** tab has settings for how the software calls bases. The basecaller you select is determined by the instrument and chemistry you are using. For further details on basecalling files and dye primer set selections, see Appendix A, "Basecallers and DyeSet/Primer Files."
- In the **Mixed Bases** tab, the Use Mixed Bases Identification box generates calls following the international standard IUB code for heterozygous positions. Mixed bases identification occurs only if secondary peak threshold is equal to or more than a specified percentage of the highest peak. You set the level according to sample type, reaction kit, and purification reaction.
- **Clear Range** is the region of the sequence that remains after excluding the low-quality or error-prone sequence at the 5′ and 3′ ends and the M13 primer sequence, if applicable. You can specify a range as a default. It is recommended that you always select Use reference trimming.
- The **Filter** tab sets the criteria for rejecting sequences if they do not meet minimum standards. Sequences not meeting the filter settings are not assembled.

Specifying the Basecall Settings

To specify the basecall settings:

1. Select **Tools** > **SeqScape Manager**.
2. Select the **Analysis Protocols** tab, then select the project in the list for which you want to change the settings.
3. Click **Properties**.
4. In the Analysis Protocol Editor, select the **Basecalling** tab to view the basecalling settings.



5. Select the appropriate basecaller dedicated to your instrument. For more information, refer to Appendix A, “Basecallers and DyeSet/Primer Files.”

6. Select the DyeSet/Primer settings (mobility files, having the .mob extension) for the instrument you are using. For more information, refer to Appendix A, “Basecallers and DyeSet/Primer Files.”

7. If you have short PCR products, you should end basecalling at the end of the PCR product. In this case, select the **At PCR Stop** check box.

8. You can also stop basecalling after a specified number of ambiguities, or Ns, or after a certain number of bases. Enter your changes to the settings.

   For more information on basecaller settings, refer to the *ABI PRISM® DNA Sequencing Analysis Software User Guide.*

9. For KB basecaller only, select to display the data as either a true or flat profile in the Processed Data area.

   True Profile – Displays data as processed traces scaled uniformly so that the average height of peaks in the region of strongest signal is about equal to a fixed value. The profile of the processed traces will be very similar to that of the raw traces.

   Flat Profile – Displays data as processed traces scaled semi-locally so that the average height of peaks in any region is about equal to fixed value. The profile of the processed traces will be flat on an intermediate scale (> about 40 bases).

10. For KB basecaller only, set the quality threshold.

    Call all bases and assign QV – Use this setting to assign a base to every position, as well as the QV.

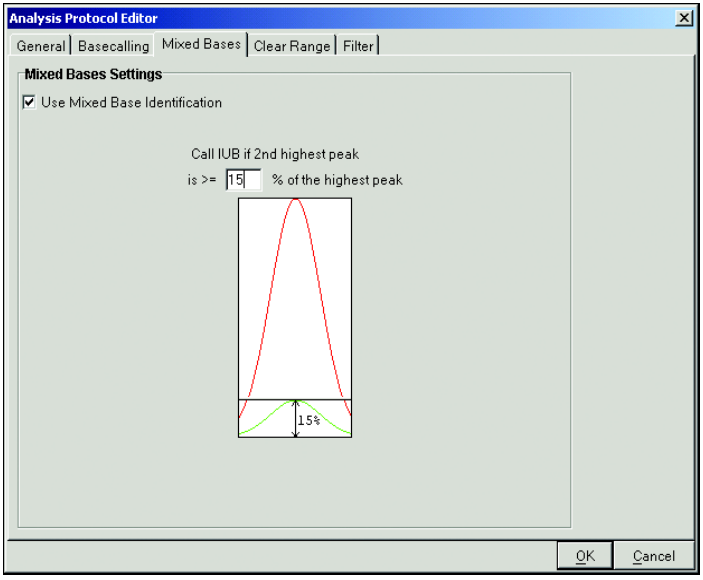
    Assign 'N' for bases with QV> – Use this setting to assign Ns to bases with QVs less than the set point. The QV will be grayed out.

IMPORTANT! **Applied Biosystems recommends that you use the KB basecaller to perform your analysis.** The KB basecaller was introduced with SeqScape software v2.0 and Sequencing Analysis software v5.0. We will continue to improve and develop this algorithm. The ABI basecaller is an older algorithm that will be removed in future versions of the software.

Specifying the Mixed Bases Settings

To specify the mixed bases settings:

1. In the Analysis Protocol Editor, select the **Mixed Bases** tab.



2. Select the **Use Mixed Base Identification** check box to generate calls according to the international standard IUB code for heterozygous positions. Mixed bases identification occurs only if the second peak height is greater than or equal to a percentage of the main peak height.

3. Set the level according to sample type, reaction kit and purification reaction, and expected or acceptable percentage. Enter the threshold for calling a mixed base for the % value of the primary peak.

   IMPORTANT! **Applied Biosystems recommends that you do not enter a value less than the 15% default value.** If you decrease the default percentage to detect low-percentage mixed bases, the background signal may be higher and interfere with mixed base detection. Be aware of this condition.

Specifying Clear Range

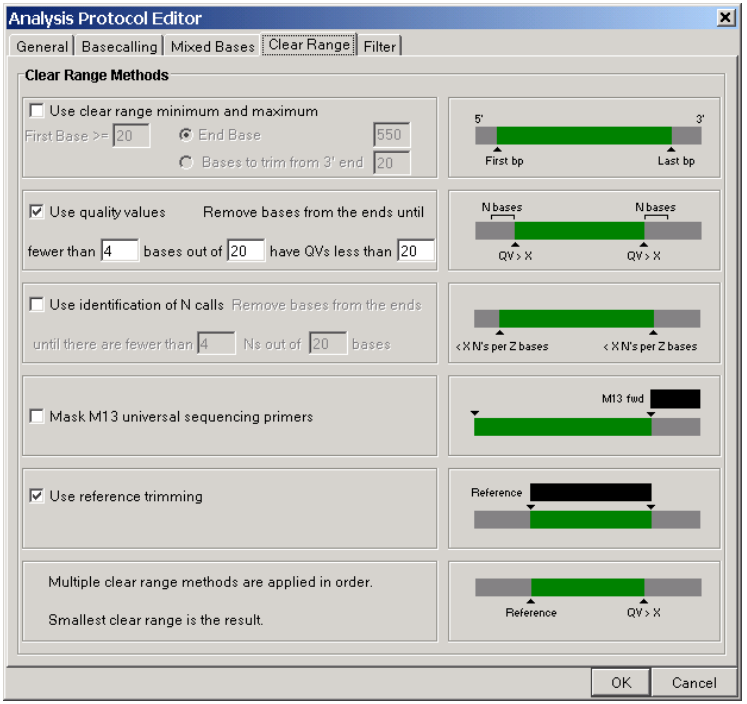
You can apply all or a subset of the Clear Range Methods algorithms. Each is applied in order from top to bottom, with the clear range method never being lengthened based on the settings in subsequent algorithms. The result is that the smallest clear range is used. If you want to preserve the existing clear range in a sample when reapplying analysis protocol settings to a sample, do not select any of the Clear Range methods.

IMPORTANT! You can create a protocol without selecting a clear range method, but it is recommended that at least one clear range method be selected for reference trimming.

The Clear Range tab enables you to set the part of the sequence that you consider to be good quality. Good quality means that the sequence has the fewest errors and ambiguities, and offers good base calling and spacing.

To set the way the clear range is determined:

1. In the Analysis Protocol Editor, select the **Clear Range** tab.



Because SeqScape software generates quality values for each base, you can choose to use a region of sequence where a certain number of bases reach an appropriate quality value.

2. Select **Use clear range minimum and maximum,** then set the minimum first base and maximum end base or the number of bases to trim from the 3’ end of the clear range.
3. Select **Use quality values** to remove bases until there are < X number of bases per Z number of bases with QV < Y. This sets a window with a specified number of allowed low-quality bases.
4. Select **Use identification of N calls** to remove bases until there are < X number of Ns per Y number of bases. This sets a window with a specified number of allowed ambiguous base calls (Ns).
5. Select **Mask M13 universal sequencing primers** to exclude the M13 primer sequence from the clear range.
6. Select **Use reference trimming** to have the samples automatically trimmed to contain only sequences that align to the reference.

Specifying the Filter Settings

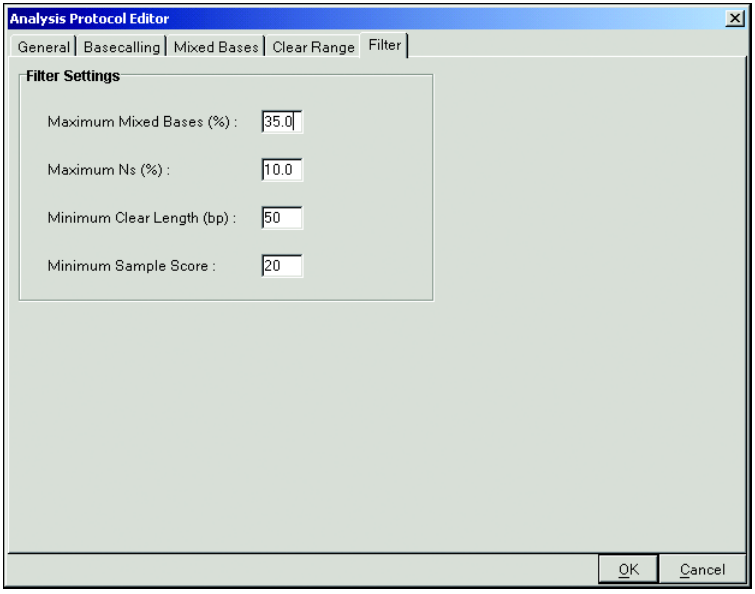
The Filter tab sets the criteria for rejecting sequences if they do not meet minimum standards. Sequences not meeting the filter settings are not assembled.

Use the maximum percentage of mixed bases to look for frame shift. Use the maximum percentage of ambiguities (N) and the minimum length settings to ensure that you are working with enough data for further analysis. This filters data that may exceed a specified percentage of ambiguities.

Also, use a minimum sample score to ensure that the quality of the sequences is high. A setting of 20 indicates that the data is accepted if the mean quality value of all bases in the clear range is 20 or greater. This corresponds to a 1-to-100, or 1%, error rate.

To select the filter settings:

1. In the Analysis Protocol Editor, select the **Filter** tab.



2. Enter your changes to the settings using the descriptions of the settings in the following table as a guide.

Parameter	Description
Maximum Mixed Bases (%)	Total maximum percentage of mixed bases that can occur in the clear range of a sample file. Any more than this number causes the sample to fail analysis.
Maximum Ns (%)	Total maximum percentage of Ns that can occur in the clear range of a sample file. Any more than this number causes the sample to fail analysis.
Minimum Clear Length (bp)	Minimum length of bases required in the clear range of a sample file. Any less than this number causes the sample to fail analysis.
Minimum Sample Score	Minimum quality value score (average of all sample QVs in the clear range) that is acceptable. The range is 1–50 (see “Sample Quality Values” on page 10-3

3. When the analysis protocol is complete, click **OK** to save the new settings. If you do not want to save the new settings, click **Cancel** to save the previous settings.

   Note: To implement the changes, you must click OK to save them and then run the analysis.

   If all filters pass, then the assembly occurs.

Specifying the Analysis Settings

To accommodate sample variability and to ensure the quality of your results, you can modify the settings used to analyze a sample and then reapply them to a sample.

You can save changes to the analysis defaults and display settings contained in a project, and you can also save them in SeqScape Manager to be used in a project template.

The procedures in the following sections describe selecting the analysis settings for a set of samples. These analysis settings can be saved as analysis defaults and saved in SeqScape Manager.

For information on reapplying a new project template, see “Reanalyzing a Project Using a Different Project Template” on page 6-22.

Gap and Extension Penalties

Project Tab Settings

The settings for Gap Penalty apply for alignment of different specimen consensus sequences to each other and the reference.

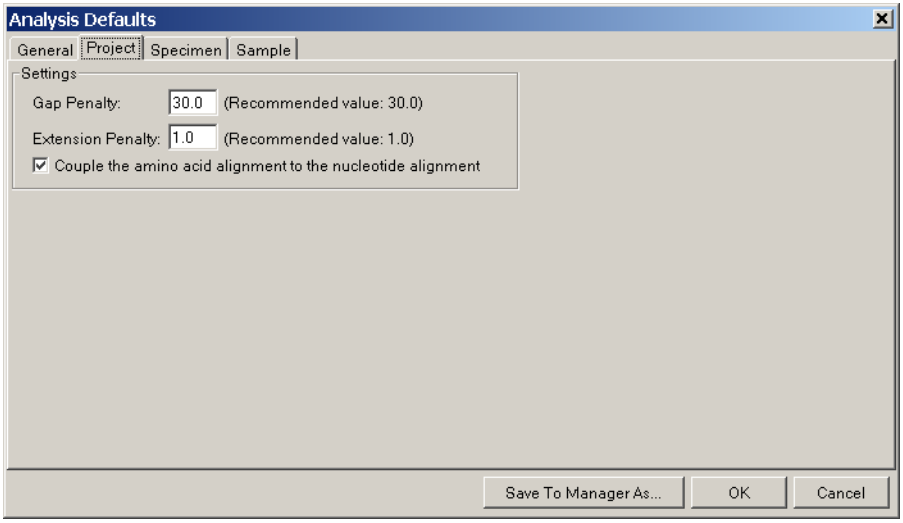


Figure 3-3 Analysis Defaults Project tab

If you want to add gap and extension penalties, these settings introduce gaps into sequence alignments allowing the alignment to be extended into regions where one sequence may have lost or gained characters not in the other gap penalty score (G+Ln). G is gap penalty, L is the length of gap, and n is the number of bases. A penalty is subtracted for each gap introduced into an alignment because gap increases uncertainty into an alignment.

Note: Applied Biosystems recommends that you use the default values. The default settings are already optimized for the current algorithm.

Specimen Tab Settings

The settings of gap and extension penalty apply to setting alignment of samples to the reference.

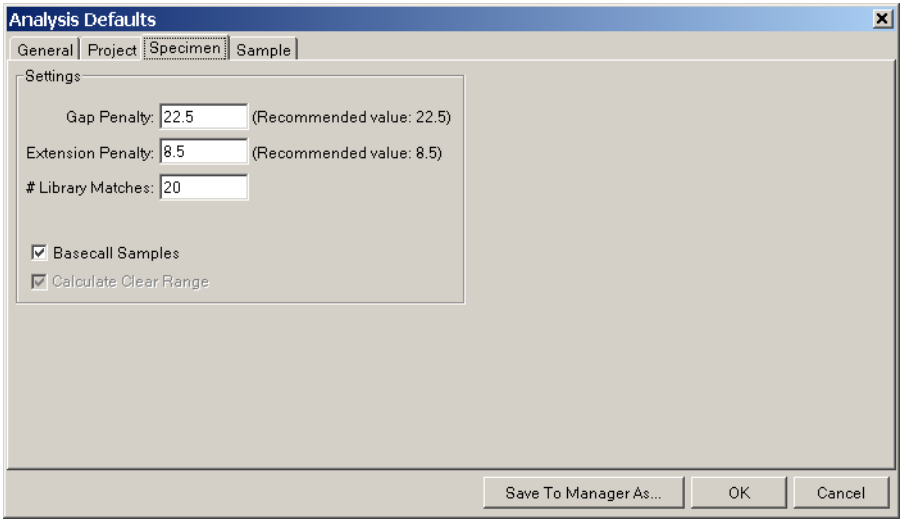


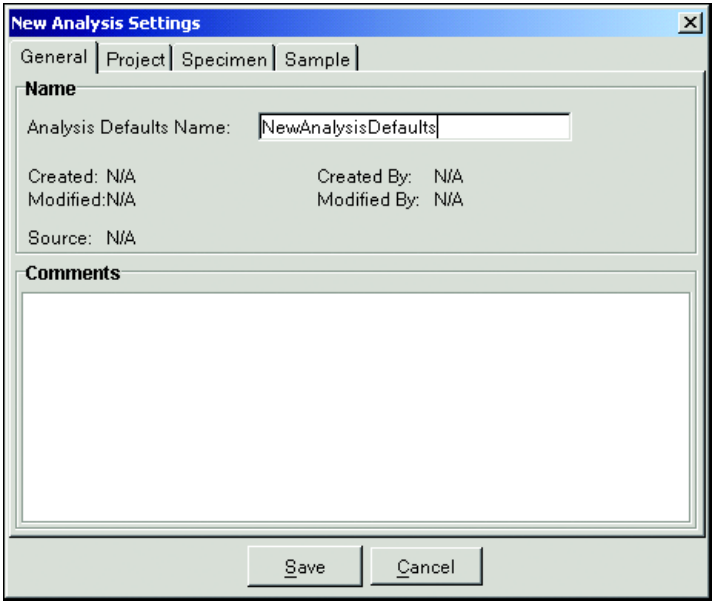
Figure 3-4 Analysis Defaults Specimen tab

Setting Analysis Defaults

To create new Analysis Defaults:

1. In the SeqScape Manager, select the **Analysis Defaults** tab, then click **New**.
2. In the General tab of the New Analysis Settings dialog box, enter an Analysis Defaults Name.

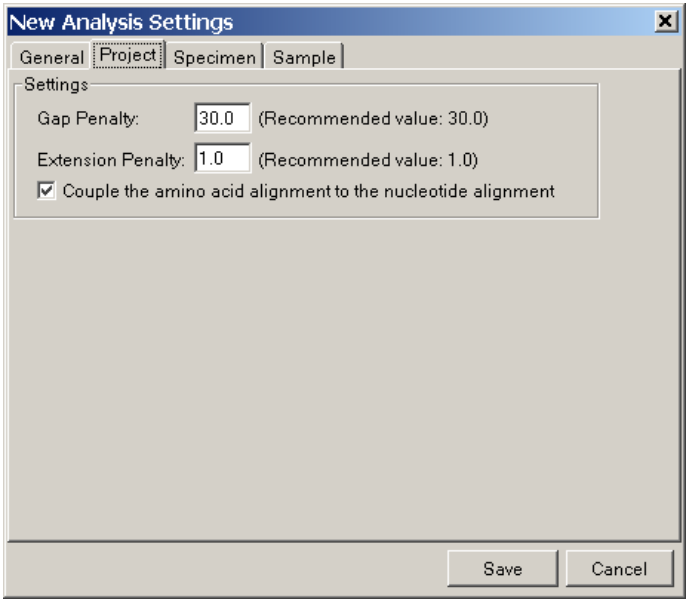
   Note: The name cannot contain spaces or characters that do not conform with the Windows file system. Refer to “File-Naming Convention” on page 2-10.



3. Enter any comments pertaining to the new analysis settings in the **Comments** box.

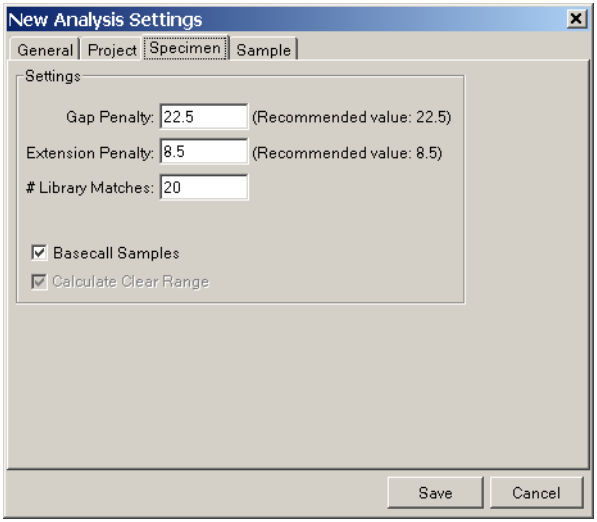
4. Select the **Project** tab and, if desired, change the Penalty Settings. The recommended gap penalty is 30 and the recommended extension penalty is 1.0.

   Note: The gap and extension penalties refer only to the alignment algorithms that are used to align the consensus sequences to the references and to each other. They do not affect the alignment of the samples to the reference for assembly.



5. Select the **Specimen** tab.

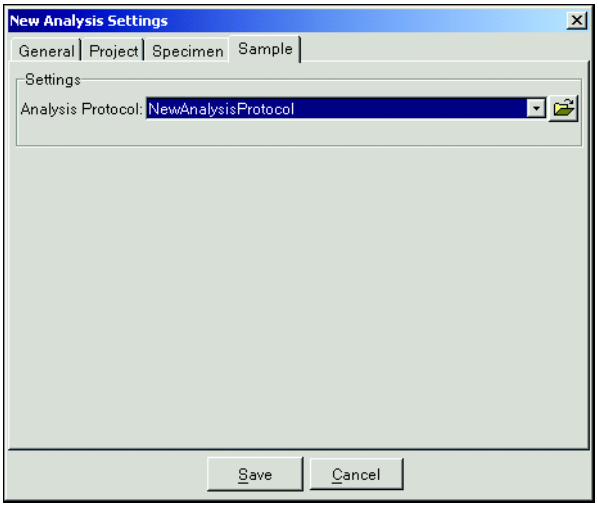
   The # Library Matches check box indicates the number of hits desired to match the library you select. The recommended gap penalty is 22.5 and the recommended extension penalty is 8.5.



   a. If desired, change the settings.

   b. Select **Basecall Samples** to automatically calculate clear range and basecall samples. If you do not select Basecall Samples, the sample files are not basecalled, and it is assumed that you have previously basecalled and edited the data. When basecalling is skipped, the software proceeds to filtering and assembly in the analysis pipeline.

6. Select the **Sample** tab then select the analysis protocol you just created from the Analysis Protocol drop-down list.



7. Click **Save** to save the new settings for this project.
8. Click **Close** in the SeqScape Manager dialog box.

Selecting the Analysis Default Settings for Individual Samples

Note: Changing the analysis defaults does not affect the analysis settings of samples that are already in the project.

To select the analysis settings for each sample individually:

1. Select the sample in the Project view.
2. Select **Analysis** > **Analysis Settings**. This opens the Analysis Protocol for that individual sample file.
3. Make relevant changes to the settings, then click **Save**.

Specifying Display Settings

To accommodate personal preferences, SeqScape software allows you to select the way results are displayed. The display settings can be modified and then reapplied to a project. The selected settings can also be saved in the SeqScape Manager to be used in a project template.

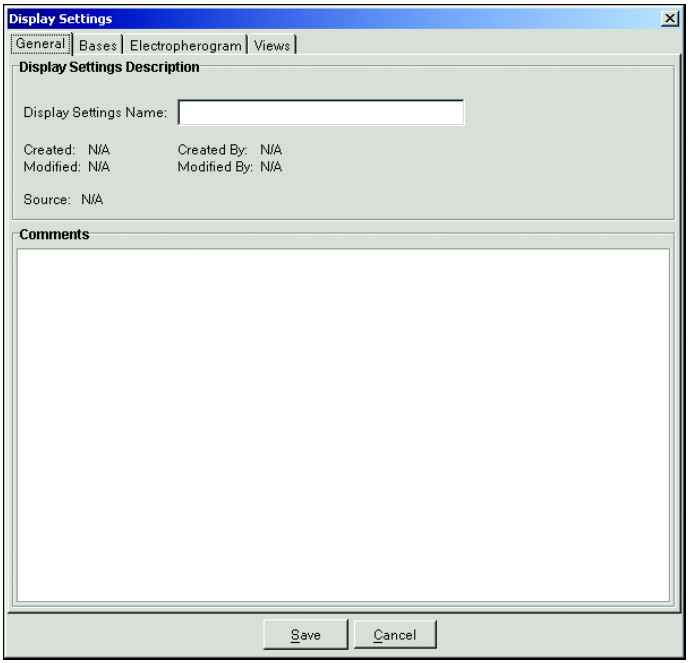
The display settings control:

- Font colors and style for bases
- Electropherogram display and axis scale
- Display views for variants
- Display views for nucleotide translation
- Quality value display and thresholds

To specify the display settings:

1. In the SeqScape Manager, select the Display Settings tab, then click **New**.

   The Display Settings dialog box opens displaying the General tab.



2. Click the **Display Settings Name** field, then enter a name for the new display settings.

   IMPORTANT! The name cannot contain spaces or characters that do not conform with the Windows file system. Refer to “File-Naming Convention” on page 2-10.

3. Enter any comments you want to record for the sample.

4. Select the **Bases** tab.

   a. Enter the desired text style and color for each base. In the Base Font section, select your font preferences for the sequence bases, or use the defaults.

   b. In the Base Scale section, enter the frequency at which to display bases for the reference sequence in the Project view.

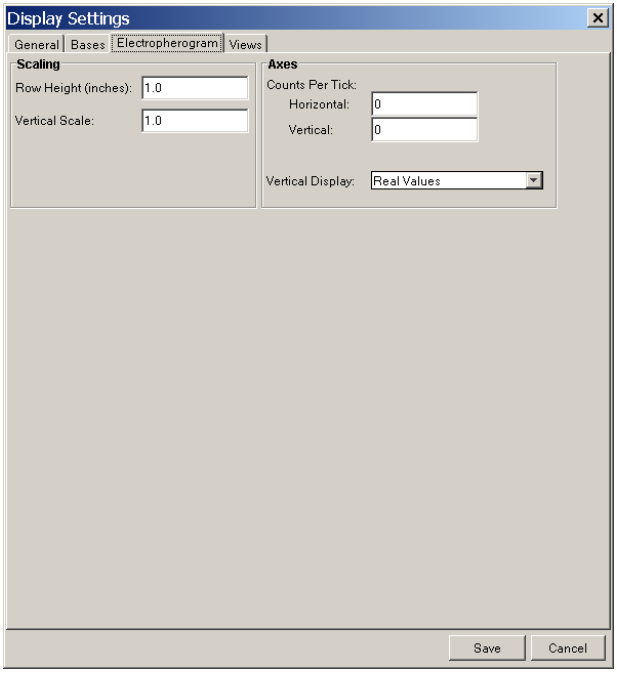
   c. In the Base Colors section, select your color preferences for the sequence bases and electropherogram traces. To select a color, click the colored box (next to A:, G:, C:, and T:) to open the color chart. Select a new color, then click **OK**.

   d. In the Quality Values section, click the colored bars to open a color chart, then select the color, if necessary. To select the threshold values, drag the divider bars between the colors.

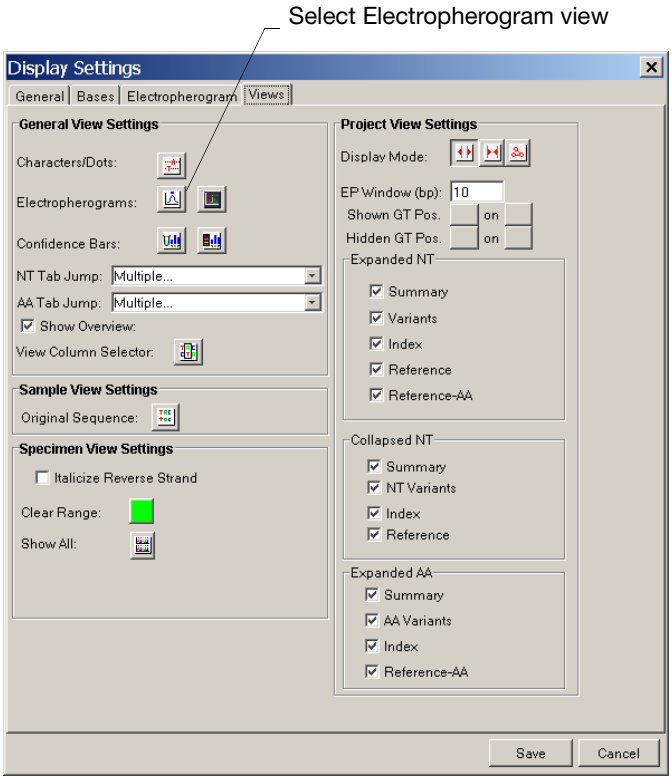
Note: The styles you specify here do not apply to variants.

For more information on quality values, see Chapter 10, “Sample and Consensus Quality Values.”

5. Select the **Electropherogram** tab, then enter your preferences for viewing the electropherogram and axes.
   a. Enter your Scaling and Axes preferences.
   b. Select a Vertical Display setting (**Real Values** or **Relative**).



6. Select the **Views** tab, then enter your preferences for the Project and Specimen views.



a. In the General View Settings section, click the buttons for the displays you want turned on in the project. Select the **Electropherogram** view.

   Note: If it is not selected, when the Assembly view is printed, it appears that the peaks are not aligned.

   Most of the buttons on this tab are the same as the viewing buttons on the lower row of the toolbar in the main SeqScape window. Refer to “Viewing Toolbar” on page 2-25.

b. In the drop-down lists, select how you want to tab through the data.

c. In the Sample View Settings section, select the icon if you want to see the original sequence displayed.

d. To differentiate forward and reverse sequences, in the Specimen View Settings section, select **Italicize Reverse Strand**.

e. In the Project View Settings section, enter the number of bases to be displayed for the Project view electropherogram snippets in the EP Window field (the minimum is 3).

7. Click **Save** to save the changes and close the dialog box.

   The new display settings are added to the SeqScape Manager.

Creating a Reference Data Group

4

This chapter contains:

Workflow for This Chapter 4-2
Reference Data Group (RDG). 4-3
Creating a New RDG Using the Wizard 4-6
Creating a New RDG Using SeqScape Manager 4-12
Defining Regions of Interest (ROI). 4-15
ROI Tab Descriptions 4-18
Creating a Library 4-20
Creating New Layers. 4-24
Declaring Variants into an RDG 4-29
Creating an RDG from Aligned Consensus Sequences 4-34

Workflow for This Chapter

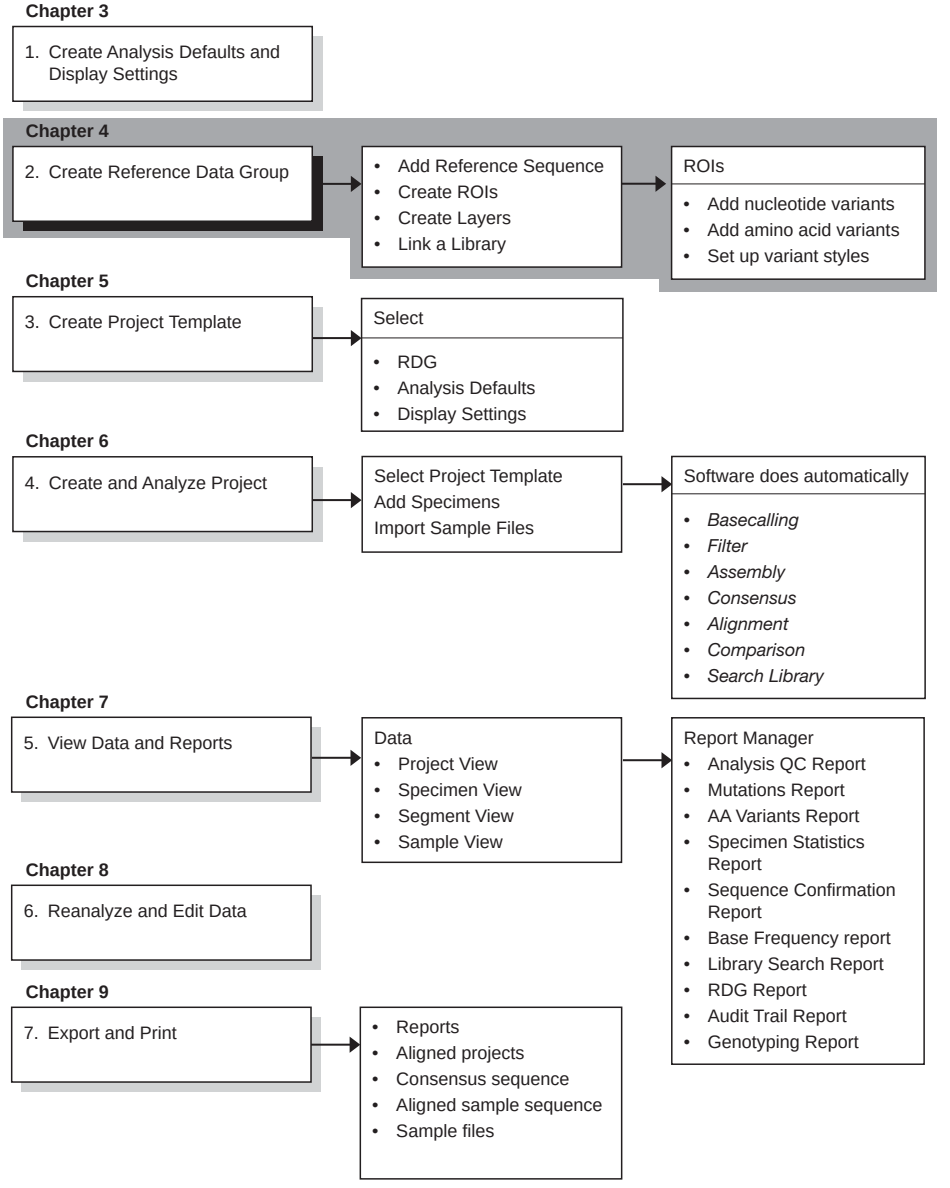


Figure 4-1 Showing Create Reference Data Group Step

Reference Data Group (RDG)

About the Reference Data Group

The Reference Data Group defines the sequence to which the ABI PRISM® SeqScape® Software Version 2.1 compares the consensus segments to the reference sequence. It contains the reference sequence and reference-associated data. The reference sequence is the entire "backbone" sequence for the project, consisting of one or more reference segments separated by reference breaks.

The RDG contains all the gene/analysis-specific information consisting of:

- A reference sequence containing continuous or discontinuous sequences made up of one or more reference segments
- Nucleotide variants
- Amino acid variants
- Translation codon table
- Layers, which are units of analysis in any project, and regions of interest (ROIs) grouped together into layers for display and translation
- Associated allele libraries
- User-defined styles for identification of variants in the project

A reference segment is a contiguous segment of the reference sequence corresponding to a single contiguous DNA sequence. It is also a region of interest. The reference segment consists of:

- An analyzed sample file
- A text-only format, FASTA, or .seq file
- A GenBank format file

GenBank Features

Every GenBank entry has a single contiguous sequence associated with it. This is also referred to as the source feature. This sequence is always numbered starting at 1.

Because of this, the sequence from a single GenBank entry translates into a single reference segment in the extended RDG. Numbering of the base ROI on this segment is set by default to start at 1.

Every GenBank entry has a feature table. These features translate into regions of interest and layers in the extended RDG. In the following table, items in {} are qualifiers read for that feature key (for example, {gene} is the value of the `\gene` qualifier). If that qualifier doesn't exist, then "" is substituted.

Table 4-1 GenBank feature table

GenBank Feature	Extended RDG equivalent
source	Skipped. The source feature corresponds to the region of interest associated with the whole reference segment that is automatically created.
exon	Region of interest is created, called {gene}_exon{number}. Translatable by default.
intron	Region of interest is created, called {gene}_intron{number}. Not translatable by default
gene	Region of interest is created, called {gene}_gene. Translatable by default.
CDS	Layer is created, called ({gene}|{product})_CDS. If translatable regions of interest exist that correspond to this CDS, then those are used for building the layer. Otherwise, new regions of interest are created as required. New ROIs are called {layerName}_region1, {layerName}_region2, etc… Translation frame and orientation is taken from CDS qualifiers (`complement()` and `\codon_start`).
misc_feature	Region of interest is created called {note}. Not translatable by default.
Unknown feature	Region of interest is created called {feature key}. Not translatable by default.

It is possible with this translation table to create many non-uniquely named ROIs (for example, if the entry had lots of variation features).

Downloading a GenBank File

To download a GenBank file from the Internet:

1. Open your web browser and enter the following URL:

   `http://www.ncbi.nlm.nih.gov/`

2. In the Search menu, select **Nucleotide**.
3. Enter the nucleotide you want for the reference sequence in the **for** text box, then click **Go**.
4. After finding the desired sequence, select the check box to the left of the accession number.
5. At the top of the page (next to Display) select **Genbank**, then select **Send to File**. The file will be saved to the specified location. This file can then be imported into the RDG. The file can have a .gb, .fcgi, or.cgi extension.

About Creating a New Reference Data Group (RDG)s

IMPORTANT! Only a user from the Administrator or Scientist group can set up a new RDG. Refer to Appendix D, "User Privileges," for a list of the privileges that apply to each group.

You can create a new RDG in the following ways:

- Use the RDG wizard and follow the instructions
- Use the SeqScape Manager window to open a blank RDG

Follow the RDG wizard procedures below, if desired, to familiarize yourself with the windows of the RDG. Then, create subsequent RDGs by using the SeqScape Manager. Refer to "Creating a New RDG Using SeqScape Manager" on page 4-12.

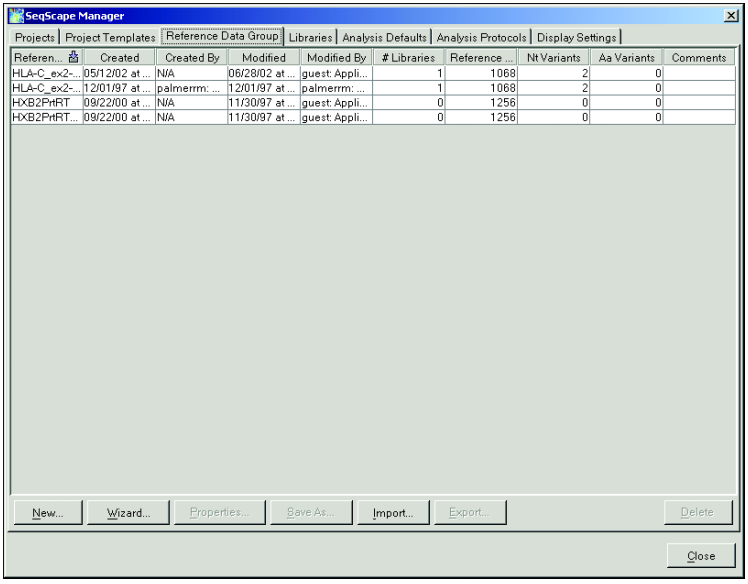
Creating a New RDG Using the Wizard

Using the Wizard to Learn the Software

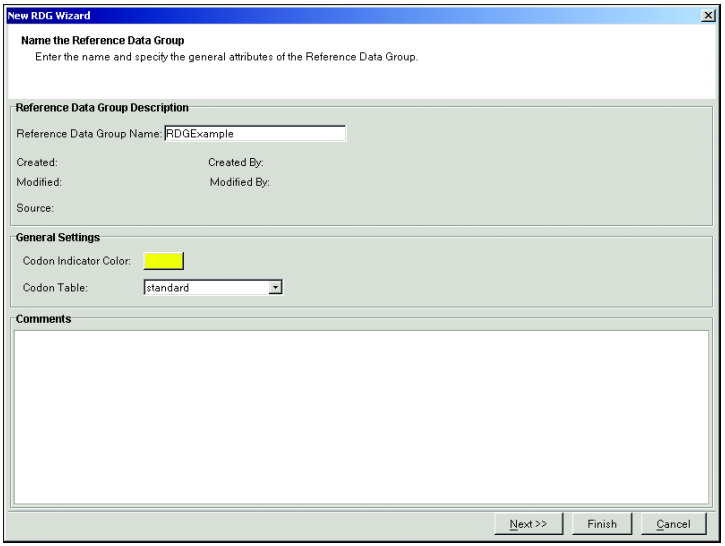
The wizard in the SeqScape Manager will familiarize you with setting up a new RDG.

To create an RDG using the RDG wizard:

1. In the main SeqScape window, select **Tools > SeqScape Manager**.
2. Select the **Reference Data Group** tab, then click **Wizard** at the bottom of the page.

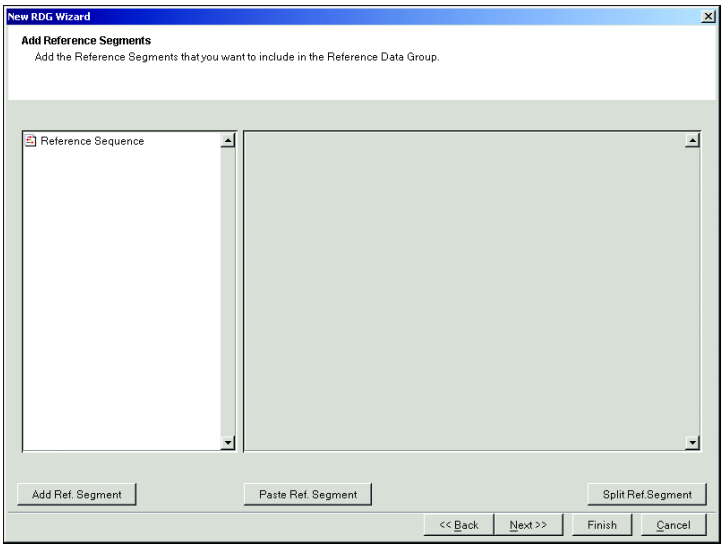


3. Enter a name for the new RDG that conforms with the Windows file system. Refer to “File-Naming Convention” on page 2-10.

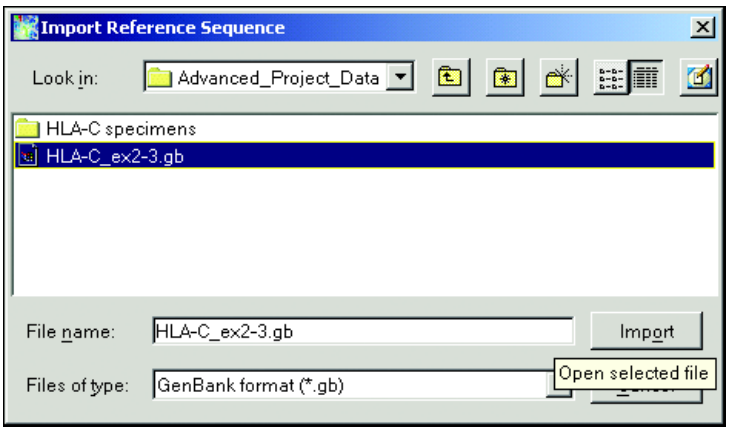


4. If desired, click the Codon Indicator Color by clicking the yellow color box, and select a new color.
5. Then select the Codon Table to use.
6. Click **Next**. The next page shows the Reference Sequence pane. The Reference Sequence forms the backbone for comparison. It is made up of one or more reference segments.

7. Click **Add Ref. Segment** in the lower left to add a segment to the Reference Sequence. A reference segment is a single sequence imported from a text file or GenBank file.

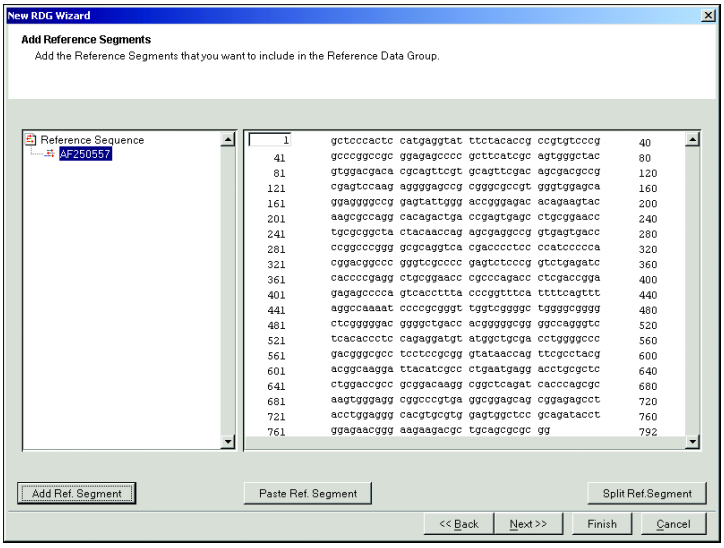


8. Navigate to the file containing the reference sequence that you have stored, such as a GenBank file (the file may have a .gb extension).



IMPORTANT! The window opens to the directory that was set up during installation of the software. If no default directory has been specified, the window opens to the C:\ drive. If you need to set up the default directory, select **Tools** > **Options** > **General**, then click **Browse** to locate the directory.

9. Click **Import**. The imported sequence appears in the right pane of the dialog box, as shown in the figure below.


New RDG Wizard
Add Reference Segments
Add the Reference Segments that you want to include in the Reference Data Group.
Reference Sequence
AF250557
1 gctcccactc catgaggtat ttctacaccg ccgtgtcccg 40
41 gcccggccgc ggagagcccc gcttcatcgc agtgggctac 80
81 gtggacgaca cgcagttcgt gcagttcgac agcgacgccg 120
121 cgagtccaag aggggagccg cgggcgccgt gggtggagca 160
161 ggaggggccg gagtattggg accgggagac acagaagtac 200
201 aagcgccagg cacagactga ccgagtgagc ctgcggaacc 240
241 tgcgcggcta ctacaaccag agcgaggccg gtgagtgacc 280
281 ccggcccggg gcgcaggtca cgacccctcc ccatccccca 320
321 cggacggccc gggtcgcccc gagtctcccg gtctgagatc 360
361 caccccgagg ctgcggaacc cgcccagacc ctcgaccgga 400
401 gagagcccca gtcacctttа cccggtttca ttttcagttt 440
441 aggccaaaat ccccgcgggt tggtcggggc tggggcgggg 480
481 ctcgggggac ggggctgacc acgggggcgg ggccagggtc 520
521 tcacaccctc cagaggatgt atggctgcga cctgggccc 560
561 gacgggcgcc tcctccgcgg gtataaccag ttcgcctacg 600
601 acggcaagga ttacatcgcc ctgaatgagg acctgcgctc 640
641 ctggaccgcc gcggacaagg cggctcagat cacccagcgc 680
681 aagtgggagg cggcccgtga ggcggagcag cggagagcct 720
721 acctggaggg cacgtgcgtg gagtggctcc gcagatacct 760
761 ggagaacggg aagaagacgc tgcagcgcgc gg 792
Add Ref. Segment
Paste Ref. Segment
Split Ref.Segment
<< Back
Next >>
Finish
Cancel

10. Click **Next**.

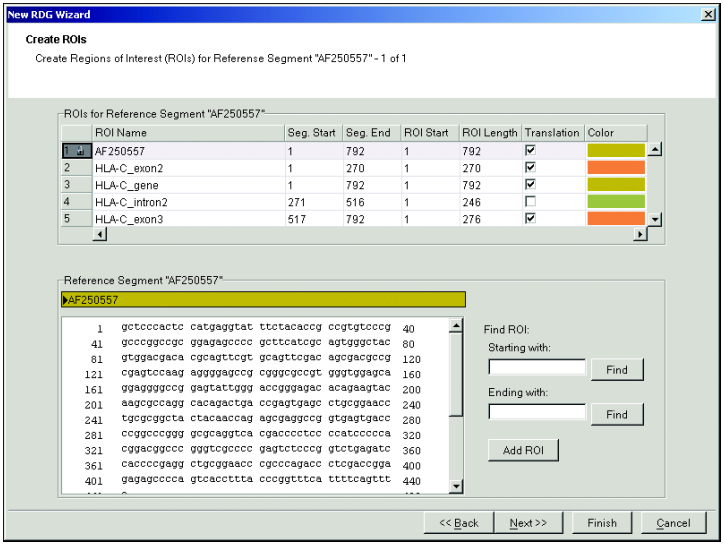
Note: For a procedure on using the Paste Ref. Segment button, refer to “Pasting a Reference Segment” on page 4-15, and for a procedure on using the Split Ref. Segment button, refer to “Adding a Reference Break in a Sequence” on page 4-27.

The wizard continues the instructions to add a new layer and regions of interest (ROI) to that layer. An ROI is a region on a reference segment that defines exons, introns, splice junctions, and other features.

Setting Up the Reference Segment

To set up the reference segment:

1. Select the bases in the region of interest that you want to compare with the reference sequence (or backbone). In the Reference Segment pane, drag through the bases you want to select, or type the starting and ending bases under the Find ROI label.

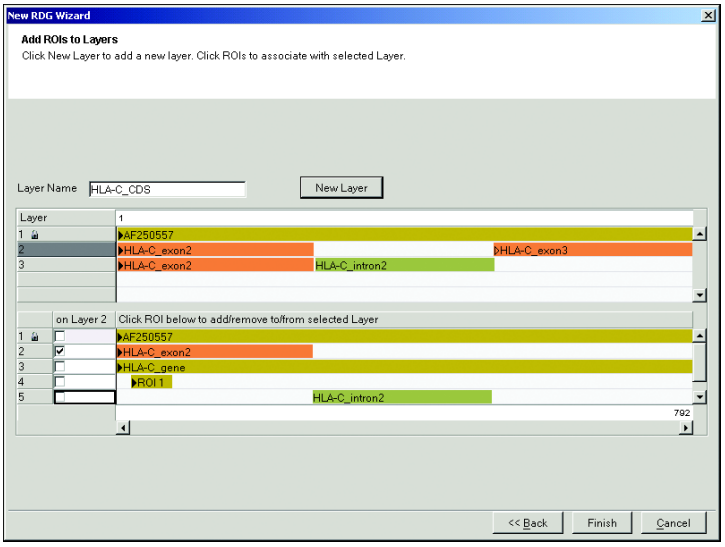

New RDG Wizard
Create ROIs
Create Regions of Interest (ROIs) for Reference Segment "AF250557" - 1 of 1
ROIs for Reference Segment "AF250557"
ROI Name | Seg. Start | Seg. End | ROI Start | ROI Length | Translation | Color
1 AF250557 1 792 1 792
2 HLA-C_exon2 1 270 1 270
3 HLA-C_gene 1 792 1 792
4 HLA-C_intron2 271 516 1 246
5 HLA-C_exon3 517 792 1 276
Reference Segment "AF250557"
AF250557
Find ROI:
Starting with:
Find
Ending with:
Find
Add ROI
<< Back
Next >>
Finish
Cancel

2. Click **Add ROI** to add the segment to the ROI table in the ROI pane above the sequence. Add as many ROIs as desired.
3. Click **Next**.
4. Follow the instructions to add layers and ROIs to layers. Layer 1 is always the reference sequence, which is generated by the software and is locked. Click **New Layer**, then name each layer that you add.

   IMPORTANT! To avoid confusion, give each layer that you add a unique name.

5. Click the new layer under the Layer label in the layer pane, then select the ROI **on Layer** check box in the ROI pane to associate it with the selected layer. Do this for each layer you create.

   IMPORTANT! In a layer, you cannot define ROIs that overlap one another.



6. Click **Finish**, or if you want to change any of the selections, click **Back**.

   The newly created RDG appears in the Reference Data Group list.

Creating a New RDG Using SeqScape Manager

Before You Begin

You must have administrator or scientist privileges to create a new RDG using SeqScape Manager.

Before creating a new RDG, make sure you:

- Download a GenBank file, a FASTA text file, or have a reference sequence that is stored on your computer
- Define on paper the ROIs, layers, and segments to compare to the reference sequence

Creating an RDG from SeqScape Manager

Creating a Reference Data Group, requires that you:

- Import reference segments
- Create ROIs
- Create layers

To create a new RDG from the SeqScape Manager:

1. In the main SeqScape window, select **Tools** > **SeqScape Manager**, then select the **Reference Data Group** tab.
2. Click **New**.
3. In the General tab, enter a name in the Reference Data Group Name field. The General tab contains general information about the RDG.
4. Select a Codon table type and add comments, if desired.
5. Select the **ROI** tab.

IMPORTANT! Do not click OK. More steps are needed to set up the RDG.

About the Reference Sequence

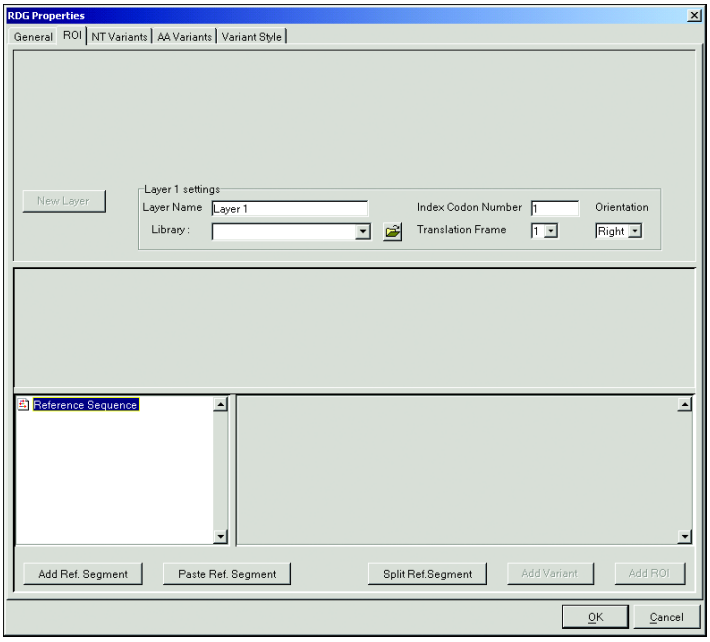
The reference sequence is made up of one or more reference segments that become a backbone or reference to which all other sequences or regions of interest are compared. After the reference sequence is imported into the RDG, it cannot be changed or edited.

Importing a Reference Segment

To form the reference sequence, you need to import one or more segments.

To import a reference segment:

1. If it is not already open, select the **ROI** tab. The dialog box that opens shows Reference Sequence as a place holder in the lower left pane.



2. Click **Add Ref. Segment** in the lower left to add a segment to the reference sequence.
3. Navigate to the file containing the reference sequence. It can be a GenBank file or a file that you stored on your computer (the file may have a .gb extension).
4. Click **Import**.

   The reference sequence is on Layer 1, which is locked so it cannot be modified.

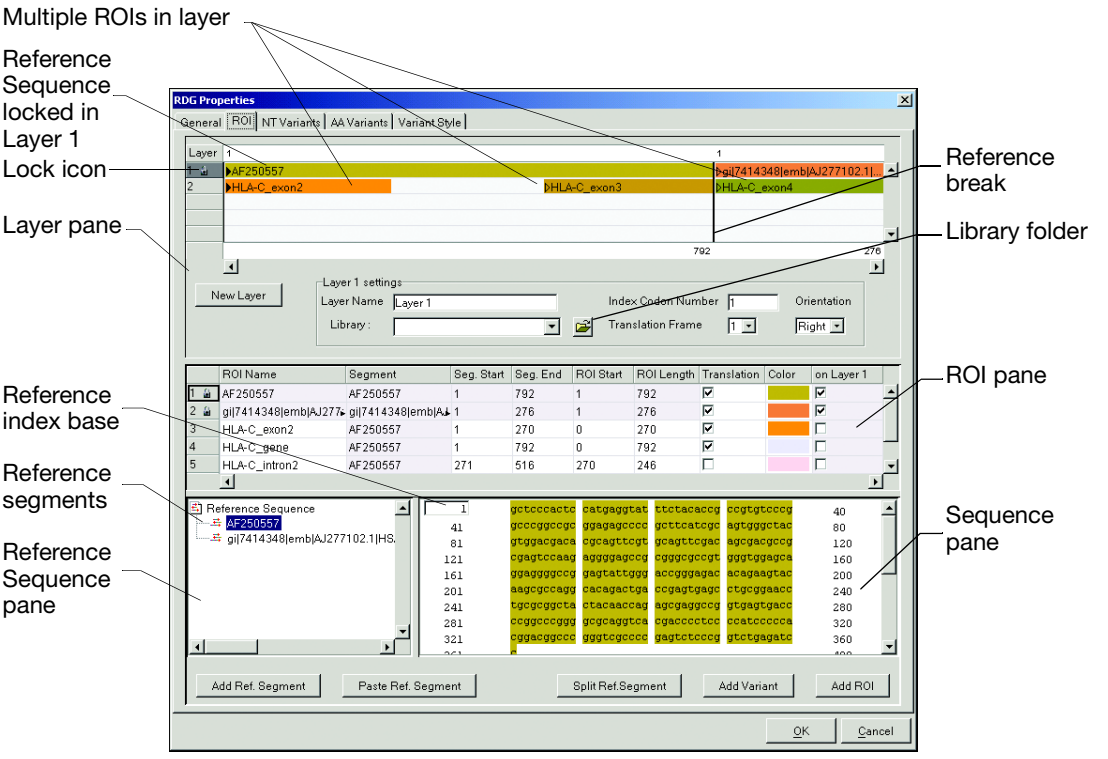


Figure 4-2 ROI Tab in the RDG Properties Dialog Box

Defining Regions of Interest (ROI)

Defining an ROI

Each reference segment has its own locked ROI. Identify the ROIs you want to define on a piece of paper, then use the information to define ROIs in the software. However, if you are using a GenBank file, the ROIs or features will already be defined. You can add additional ROIs where appropriate to your analysis.

To define an ROI:

1. In the ROI tab, select an empty layer or a layer where you want the ROI to appear.
2. Select a segment in the nucleotide sequence pane (by dragging through the region of interest), then click **Add ROI**.
3. Enter a name for the ROI under the ROI Name column in the ROI pane.
4. Define as many ROIs as appropriate by dragging through the regions of interest, or by entering a number in the text box where the ROI should begin.

Pasting a Reference Segment

You can create or enter a sequence in a text editor or word processing program and copy the segment into the RDG at a later time.

To define a reference segment by pasting:

1. Open a text file, then click-drag the region of interest you want to use as a reference segment.
2. Select **Edit** > **Copy**.
3. In the RDG Properties ROI tab, click **Paste Ref. Segment** to use a reference segment that you copied to the clipboard. The copied reference appears in the Reference Sequence pane.
4. If you want to delete the copied reference segment, select it, press **Delete**, then click **OK** in the Confirmation dialog box.

Deleting an ROI or Layer

To delete an ROI or layer, reference segment:

1. Select the ROI, layer, or segment.
2. Press **Delete**. Only unlocked rows can be deleted.
3. Check this on software

   IMPORTANT! After you delete an object, it cannot be undone.

Deleting a Reference Segment

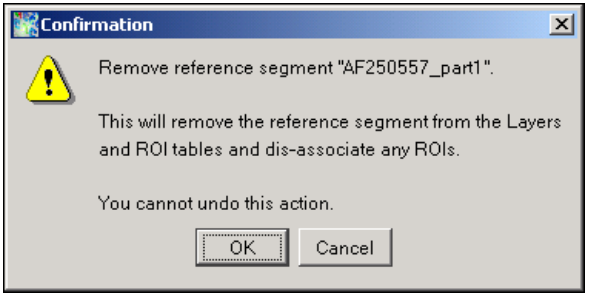
When RDG Properties window is open without being associated with a project, a reference segment can be deleted. Right-clicking a selected reference segment in the Reference Sequence pane opens a pop-up menu with selections to rename or delete the selected segment.

Note: The reference segments and Layer 1 cannot be deleted by selecting them in the Layer pane, then pressing Delete, because they are in locked layers. They cannot be deleted when they are part of an existing open project.

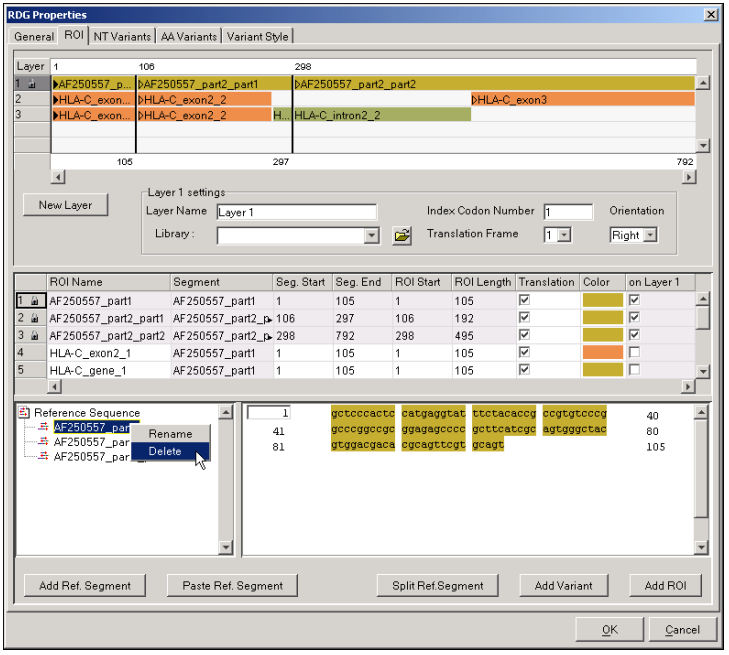
To delete a reference segment:

1. In SeqScape Manager, select the **Reference Data Group** tab and highlight the RDG in which you want to delete a reference segment.
2. Click **Properties**, then click the **ROI** tab.
3. In the Reference Sequence pane, highlight the reference segment you want to delete.

4. Right-click and select **Delete** in the pop-up menu.
   A confirmation dialog box opens.


Confirmation
Remove reference segment "AF250557_part1".
This will remove the reference segment from the Layers and ROI tables and dis-associate any ROIs.
You cannot undo this action.
OK
Cancel

5. Click **OK** in the confirmation dialog box. Once you click OK, the delete cannot be undone.



ROI Tab Descriptions

Layer Pane Functions

The Layer pane in the ROI tab (refer to Figure 4-2 on page 4-14) has the following functions:

- Layers – Shows the locked Reference Sequence in Layer 1, and ROIs associated with each layer.
- New Layer button – Adds a new layer to the end of the layer table.
- Layer Number Settings – The settings of the selected layer. Each layer has its own unique settings.
- Layer Name – The name of the layer, which can be edited.
- Library – Contains libraries to select if you are performing allele or haplotype identification. Before you select a library to associate, the Library field is blank. A library can be copied into the RDG, but is not associated until you select it from the Library drop-down list. Once it is selected, the Library field shows the name of the library.
- Index Codon Number – Indicates the first amino acid number. This number is always in relation to the number of the first reference segment base, positive numbers only.
- Translation Frame – Sets the translation frame for the layer. The values are 1, 2, 3.
- Orientation – Sets the orientation of the layer, right (forward) or left (reverse).

The ROI Pane

The ROI pane has the following features:

- Clicking a row selects the ROI. When you select an ROI in the RDG, it selects and scrolls the reference segment and the associated sequence.
- Primary ROIs that are created when reference segments are imported are locked as indicated by the lock icon. These primary ROIs cannot be deleted from the ROI table, but can be deleted from the Reference Sequence navigation pane by right-clicking and selecting Delete.

Columns in the ROI Pane

The ROI pane in the middle of the RDG Properties dialog box has the following columns:

- **ROI numbers** – The number of the ROI. The Reference Sequence on Layer 1 is always locked. Reference segments that make up the Reference Sequence are also locked. Unlocked layers are below the reference segments and can be edited.
- **ROI Name** – Name of the ROI. ROI names that are not locked can be edited. The ROI Name must be unique.

  Note: Names for Reference Segments are not editable in the ROI pane. They can be edited in the Reference Sequence navigation pane by right-clicking and selecting **Rename**.

- **Segment** – Name of the segment to which the ROI is associated.
- **Seg. Start** – The nucleotide number in the Reference Sequence where the ROI begins.
- **Seg. End** – End of the ROI segment.
- **ROI Start** – The first nucleotide number you assign to this ROI. The number can be positive or negative.
- **ROI Length** – Length of the ROI. The value is automatically recalculated if you change the Segment Start or ROI Length values. Entering a number into this cell automatically recalculates the ROI Length value.
- **Translation** – Sets whether or not the ROI is translated.
- **Color** – Shows the color of the ROI. Click to display the standard Color Picker dialog box if you want to select a different color for the ROI.

  Note: When an ROI is defined, a default color is applied to the ROI based on the name of the ROI.

- **On Layer (number)** – Check box. The label for this column changes based on the selected layer. If the check box is selected, the ROI appears on the selected layer. ROIs can be associated with multiple layers. However, ROIs cannot overlap on a layer. Therefore, the check box is disabled if the Start/Length range of the ROI overlaps with the range of an ROI already associated with the layer. This prevents you from overlapping ROIs on the Layer table. A dialog box appears if you try to select an overlapping ROI.

Creating a Library

About the Library

You must classify your library as a haploid or diploid library and determine how many library matches you would like to see for each consensus sequence. A library match is one allele or a pair of alleles that agree closely with each consensus sequence.

A haploid library contains sequences that have pure bases only (AGCT). When searching against a haploid library, SeqScape software provides library matches, and each library match contains a pair of sequences (haplotypes) that best match the genotype of each consensus sequence.

A diploid library contains sequences composed of pure bases only, or pure bases and mixed bases. When searching against a diploid library, SeqScape software provides library matches, and each match is a single sequence that best matches the genotype of each consensus sequence.

Using Aligned FASTA Files

To use the library search feature, you must import an aligned multiple sequence FASTA file into the SeqScape software. If you have a series of FASTA sequences, you must use a tool to align those sequences and create a single aligned multiple sequence FASTA file before importing the file into the software.

IMPORTANT! All sequences in the library must be of equal length. If some sequences are shorter than other sequences, you can use dashes (-) for missing bases.

Using a Tool to Align the Files

If you have a series of text sequences or electropherograms, you must create FASTA files, then use a tool to align those sequences and create a single multi-aligned FASTA file.

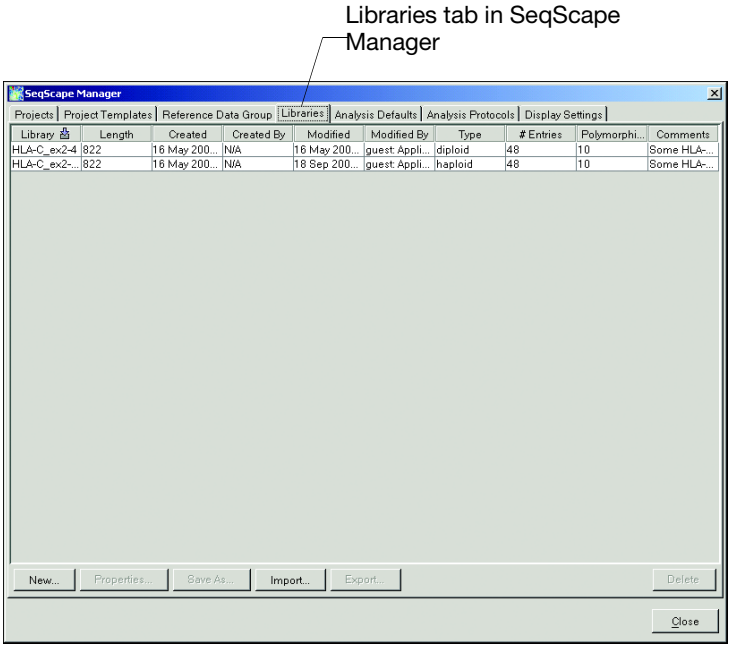
A common tool used to create aligned multiple-sequence FASTA files is Clustal X. For instructions on how to obtain and use Clustal X, see Appendix F, “Creating a Multi-Aligned FASTA File.”

Setting Up Your Library

Note: Review the following procedure, "Setting Up Your Library," then go to the next section, "Creating New Layers" on page 4-24. Use the following procedure to select the library before continuing with the procedure to create new layers.

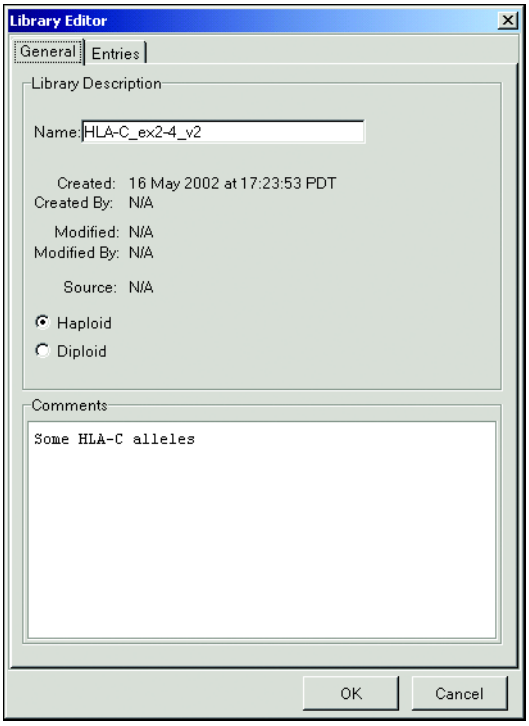
To set up a library:

1. In the main SeqScape window, select **Tools** > **SeqScape Manager**, then select the **Libraries** tab.



2. Select **New**.

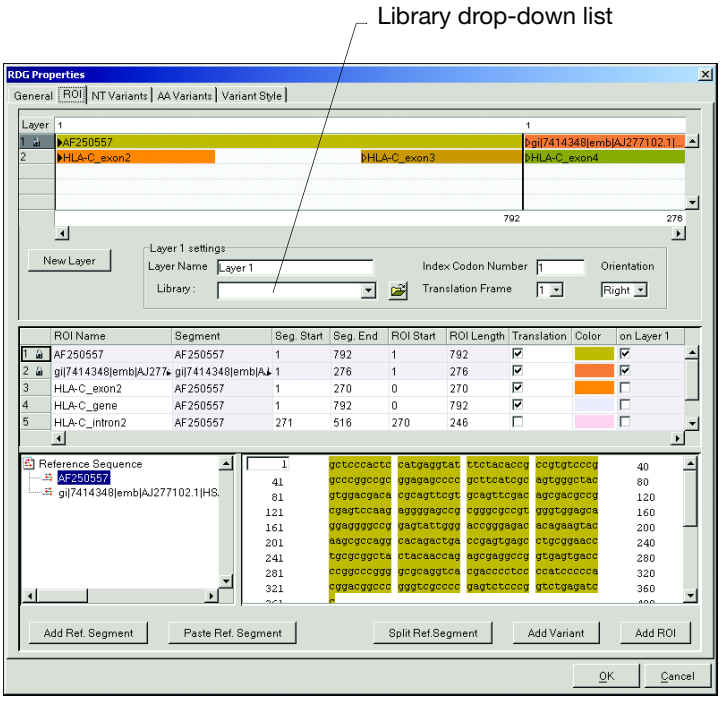
3. In the Library Editor General tab, enter a name for the new library, and select **Haploid** or **Diploid**.



4. Select the **Entries** tab, then click **Import**.
5. Import the aligned multiple sequence FASTA file, and click **OK**.
6. In SeqScape Manager, select the **Reference Data Group** tab and select the RDG that you want to link to the library
7. Click **Properties**, then select the **ROI** tab.

   Note: At this point, if you do not have layers in the RDG or you do not know how to create a layer, go to "Creating New Layers" on page 4-24. Otherwise, continue to the next step.

8. Select a layer in the Layer pane.
9. In the Library drop-down list, select the corresponding library that you created in steps 2 through 5.

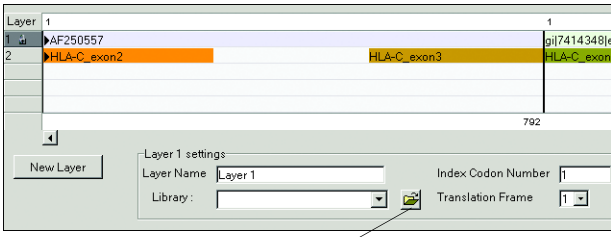


Creating New Layers

Layers organize groups of related, nonoverlapping ROIs. By organizing ROIs into layers, results reviewing and library searching are faster and more focused. The Layer table shows the organization of ROIs into layers.

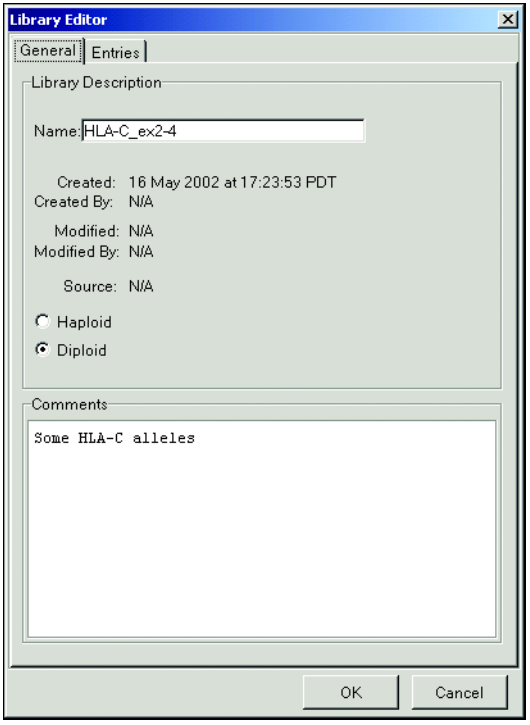
To create new layers:

1. In the ROI tab, Layer pane, click **New Layer**, then enter a name in the Layer Name field.
2. Select a layer by clicking it under the Layer label in the Layer pane. If you need more information on libraries, refer to “Creating a Library” on page 4-20.
   a. Select a library from the Library drop-down list if you are performing allele or haplotype identification.
   b. Select the library folder icon to open the Library Editor and view the entries.



Note: The selected library in the Library drop-down list is associated with the layer in the Layer Name field.

The Library Editor opens as shown in the sample below.



c. In the Layer pane, enter the Index Codon Number.

d. Select the Translation Frame.

e. Select the Orientation.

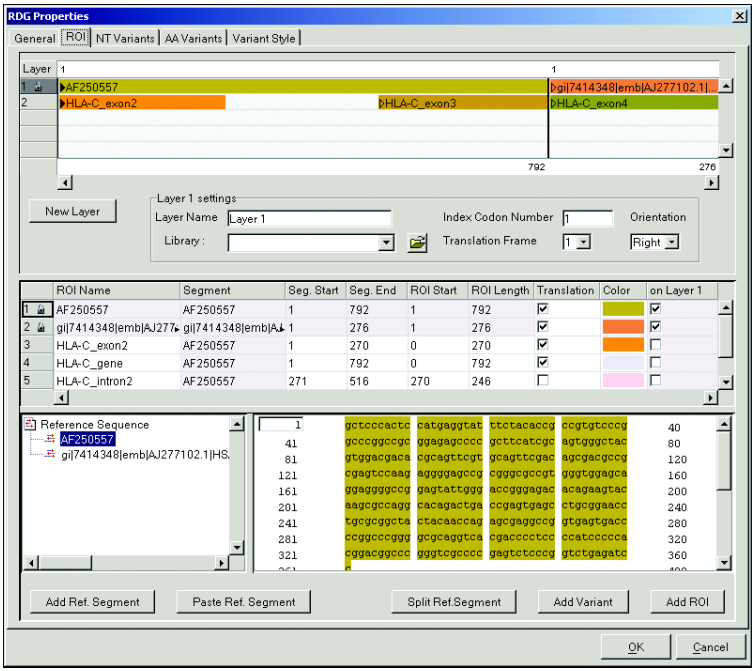
3. Select the appropriate Reference Segment in the Reference Sequence pane, highlight the sequence representing the ROI, then click **Add ROI**.

   The ROI is added to the ROI pane and to the selected layer.

4. Repeat the process to build layers containing all the ROIs and layers you previously defined on paper.

5. You can edit the ROIs in the ROI pane by selecting the attributes, then editing them directly in the table.

6. To include an existing ROI on an unlocked layer, select or create the layer, then select the **On Layer** check box for the ROI.



IMPORTANT! If you want to add variants, follow the procedure on page 4-29. Do not click OK. If you do not want to add variants, go to the next step.

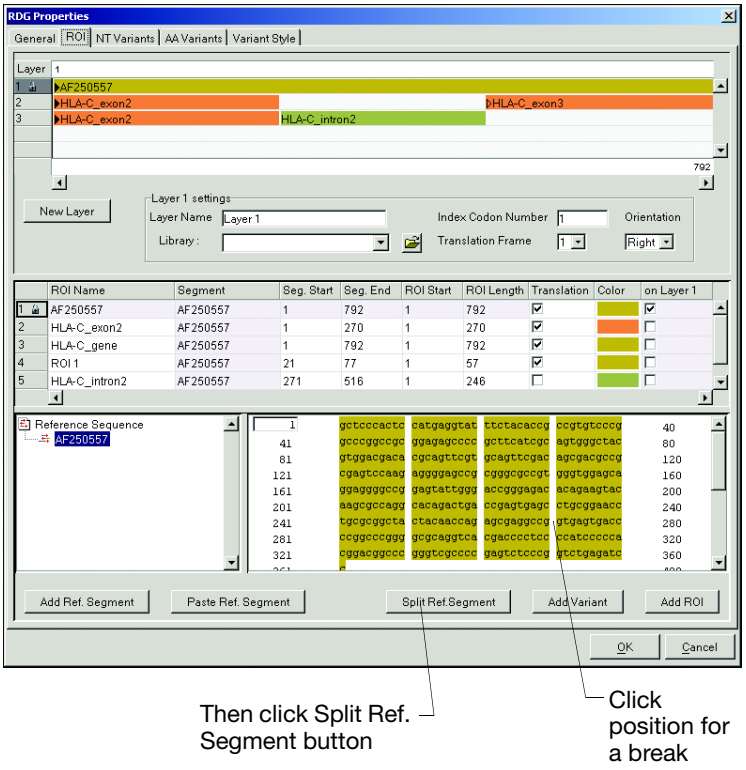
7. When you finish adding ROIs and layers, click **OK**. The new RDG appears in the Reference Data Group list.

Adding a Reference Break in a Sequence

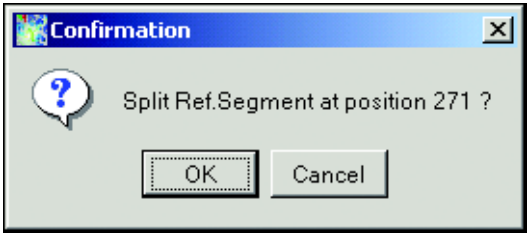
A reference break can be added in the Reference Sequence by using the Split Ref. Segment button. Reference segments can be split if you want to delete intervening reference sequences. When reference segments are split, the ROIs associated with the reference segment are also split.

To add a reference break:

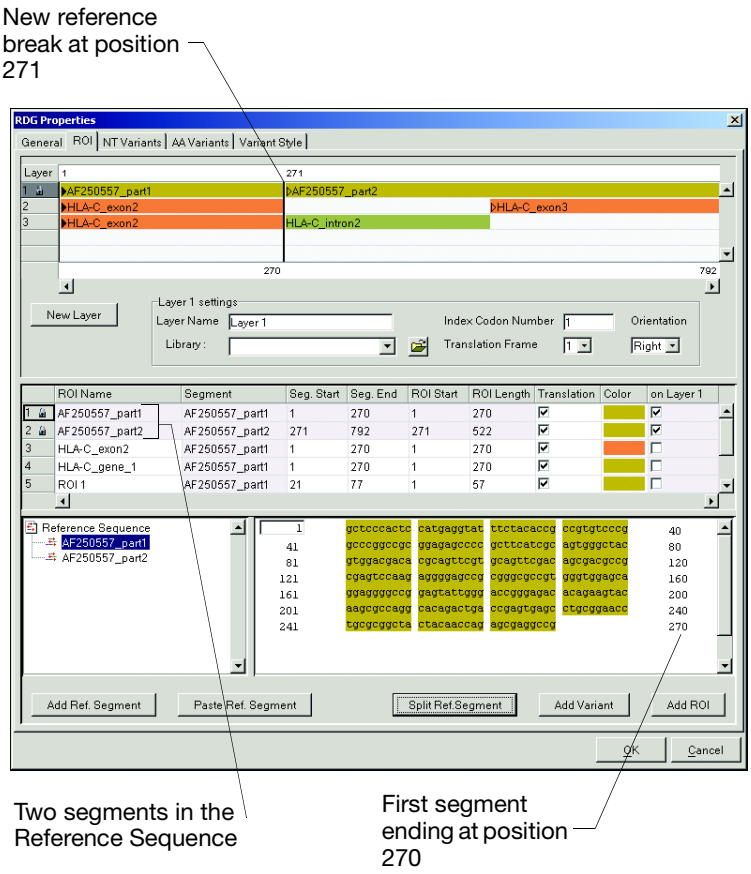
1. In the ROI tab Sequence pane, select the base position where you want a split to occur, then click **Split Ref. Segment**.



2. A confirmation dialog box opens showing the position for the reference break. Click **OK**.



There is a new reference break in the Reference Sequence in the Layer pane, as shown in the next screen shot. The Reference Sequence shown in the ROI Layer is now in two locked layers, one segment ending at position 270, and the second segment starting at position 271. The Sequence pane shows the first segment ending at position 270.



Declaring Variants into an RDG

About NT Variants

The NT Variants tab in the RDG Properties dialog box lists the known nucleotide variants associated with a reference sequence. The entries you define in this tab are used to identify known and unknown variants in your projects.

You can enter NT variants by:

- Clicking Add Variant in the ROI tab, then entering the variant attributes in the New NT Variant dialog box.
- Creating a table of variants in a tab-delimited format, then saving the file and importing it into the NT variant file.

  One way of creating a table of variants is by using Microsoft® Excel. The columns in the Excel table must map to the columns in the NT Variants tab as shown below.

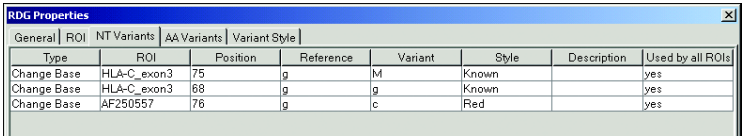


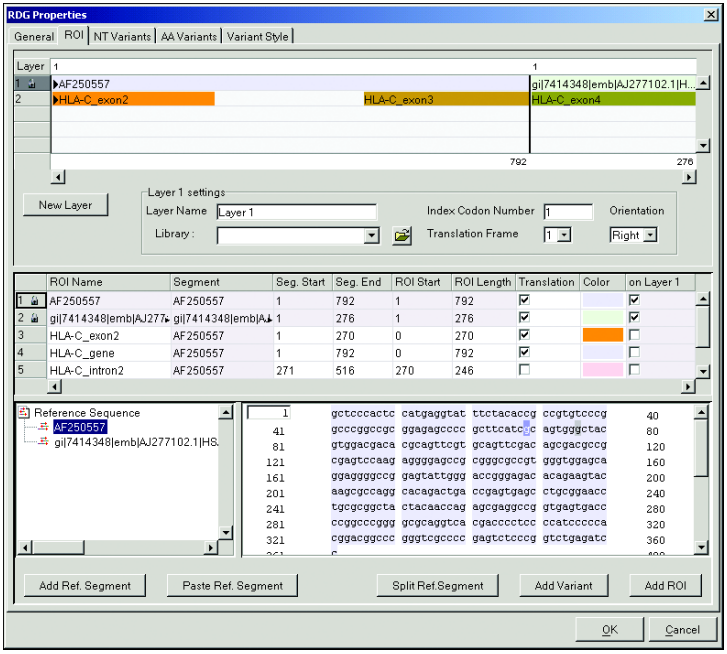
Figure 4-3 NT Variants Tab Showing Table Column Names

- Importing an aligned FASTA file.
- Selecting a sequence within a reference segment, then clicking **Add Variant**. This procedure is described below.

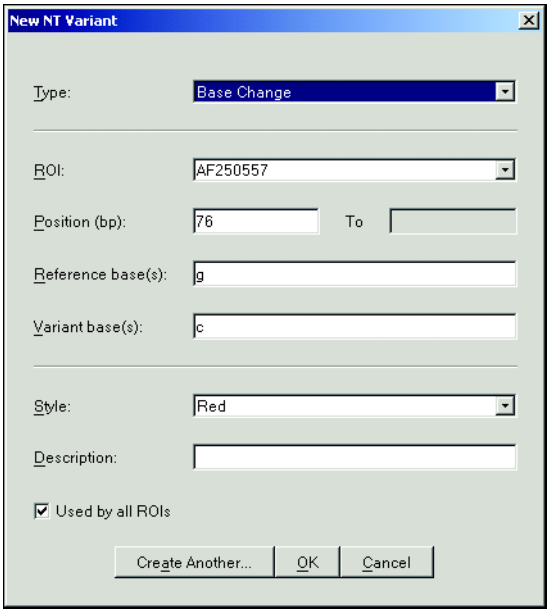
Creating New NT Variants

To create new NT variants:

1. In the SeqScape Manager, select the **Reference Data Group** tab, then click **Properties**.
2. Select the **ROI** tab.
3. Drag to select a sequence in the nucleotide sequence area of the tab, then click **Add Variant**.



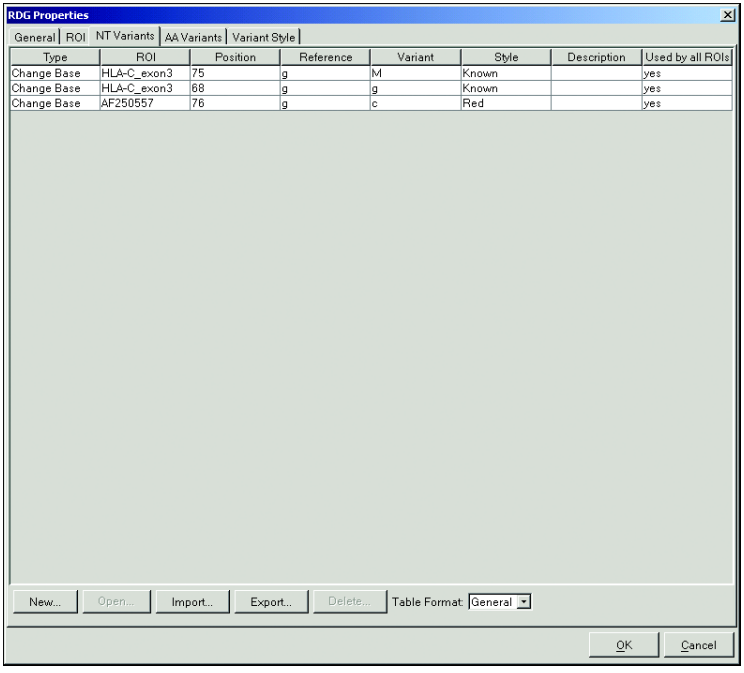
4. In the New NT Variant dialog box, select the type of variant: **Insertion**, **Deletion**, or **Base Change**.



5. Enter the Variant base.
6. If desired, change the style and enter a description.
7. Select the **Used by all ROIs** box if this NT variant is to be used by all ROIs.

8. Click **Create Another**, or **OK** to save the changes.

   After you click OK, the variant additions appear in the list in the NT Variants tab.

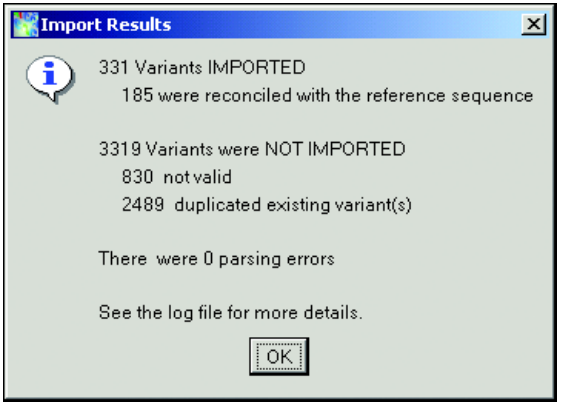


Importing NT Variants in Tab-Delimited Format

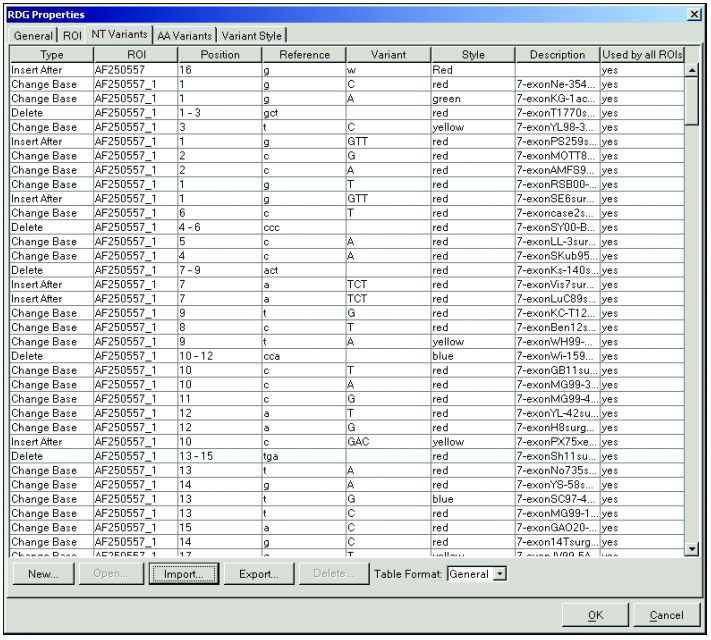
To import an NT variant from a tab-delimited NT variant file:

1. In the SeqScape Manager, click the **Reference Data Group** tab.
2. Select the RDG in the list for which you want to import NT variants.
3. Click **Properties**, then select the **NT Variants** tab.
4. Click **Import**.
5. Navigate to the tab-delimited NT variants file, and click **OK**.

6. An Import Results dialog box opens to show the number of variants imported as shown in the sample below. Click **OK** to close the Import Results dialog box.


Import Results
331 Variants IMPORTED
185 were reconciled with the reference sequence
3319 Variants were NOT IMPORTED
830 not valid
2489 duplicated existing variant(s)
There were 0 parsing errors
See the log file for more details.
OK

7. The new variants appear in the NT Variants list. The Table Format options at the bottom of the window are General (default) and Hugo. If desired, select the format in the drop-down list.



8. Click **OK** to save the imported variants and close the RDG Properties window.

Creating an RDG from Aligned Consensus Sequences

About Creating an RDG

SeqScape software will create a new reference sequence and variants from a set of aligned sequences imported into a blank RDG that contains no reference sequence. The file format of the imported aligned sequences must be in FASTA text. For more information on FASTA format, see Appendix E, “Aligned Variant and FASTA File Format.”

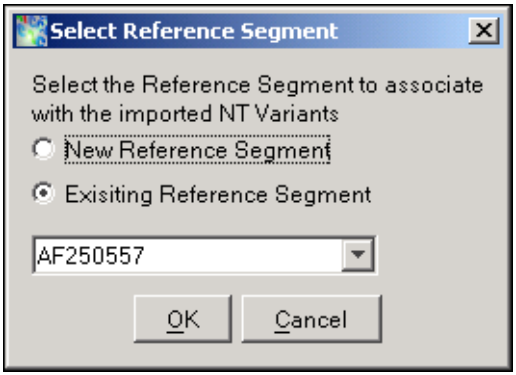
SeqScape software uses the first sequence in the set of aligned sequences in the FASTA file as the reference. The rest of the sequences will be evaluated relative to that first sequence to derive variants. Any positions that are found to differ from the first sequence will be used to populate the variants table.

Importing NT Variants from an Aligned FASTA File

To import NT variants using an aligned FASTA file:

1. Select **Tools** > **SeqScape Manager**.
2. Click the **Reference Data Group** tab, then select the RDG for which the variant will be added.
3. Click **Properties**.
4. In the RDG Properties window, select the **NT Variants** tab.
5. Click **Import**.
6. In the Import NT Variants dialog box, navigate to and select an aligned sequence FASTA file (.fsta extension).
7. Click **Import**.

8. In the Select Reference Segment dialog box, select the reference segment for which the variants are to be added.

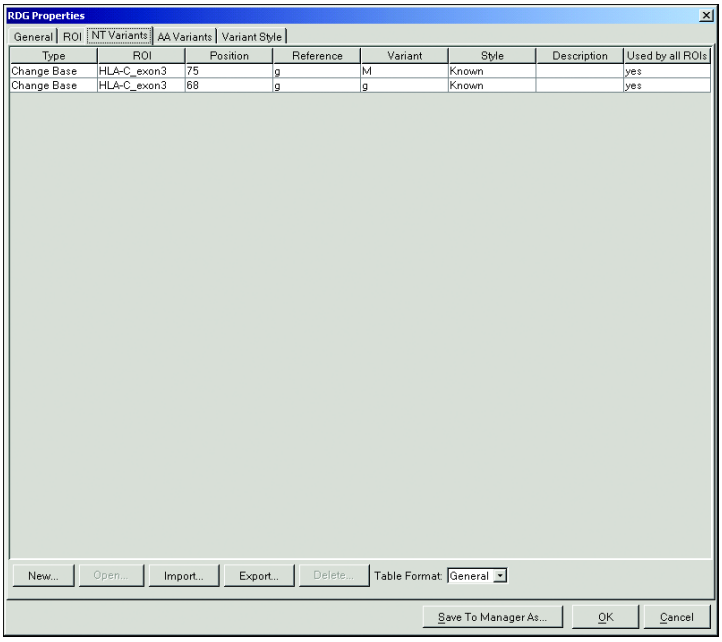


9. Click **OK**.

   After the data is imported, the Import Results dialog box opens, displaying information about the import.

   The first sequence in the imported file will populate the reference. The subsequent sequences will be used to derive variants by comparison to the first sequence. These variants will appear in the Variants table.

10. Click **OK** to close the Import Results dialog box. The list of variants are displayed in the NT Variants tab.



11. Click **OK** to close the RDG Properties window.

Entering New AA Variants

The AA Variants tab lists the known amino acid variants associated with a reference sequence. The entries you define in this tab are used to identify known and unknown amino acid variants in your projects.

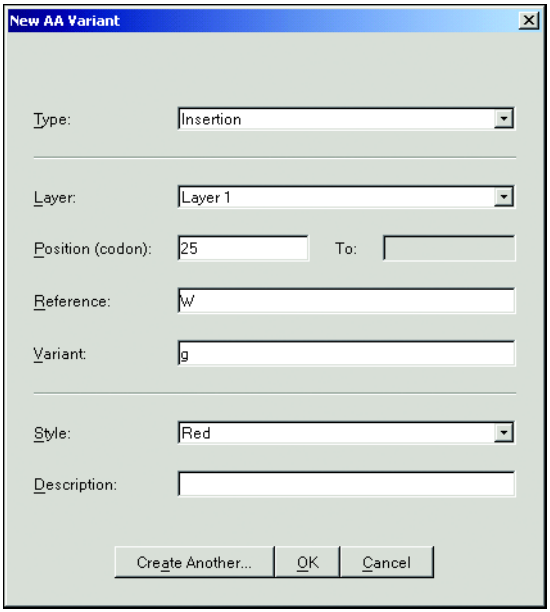
You can enter AA variants in two ways:

- Click Add Variant in the ROI tab, then enter the variant attributes in the New AA Variant dialog box.
- Create a table of variants using Microsoft® Excel, then import the table. The columns in the Excel table must map to the columns in the AA Variants tab. Refer to Figure 4-3 on page 4-29 for the column names.

To enter a new AA variant:

1. In SeqScape Manager, click the **Reference Data Group** tab.
2. Select a listed RDG, and click **Properties**.
3. Select the **AA Variants** tab, then click **New**.

4. Select the type of variant (**Insertion**, **Deletion**, or **Residue Change**).

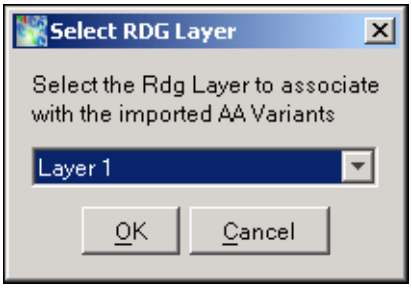


5. Enter the Position (codon) in the reference sequence that you want changed. The Reference appears after you enter the position in the sequence.
6. Enter the variant.
7. Select a color style and enter a description, if desired.
8. Click **OK**. The new variant appears in the AA Variants list.
9. Click **OK** to save the new variant.

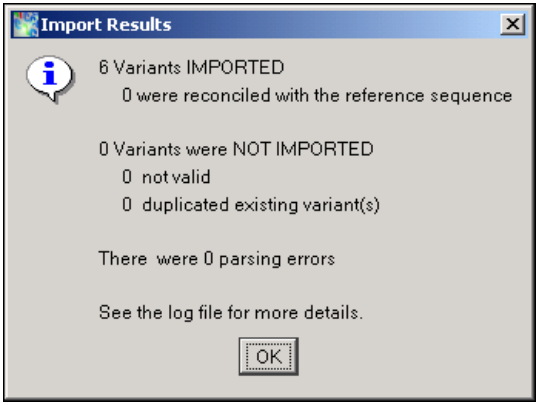
Importing AA Variants

To import an AA variant from a tab-delimited file:

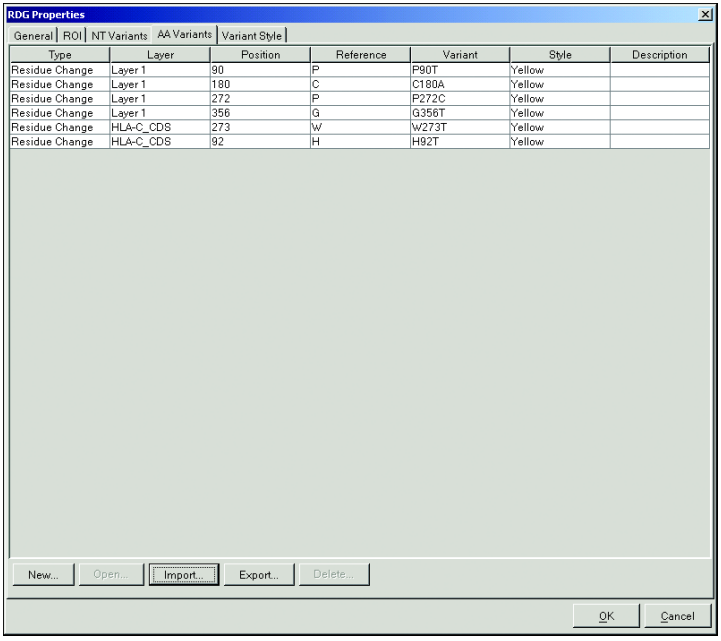
1. In SeqScape Manager, select the **Reference Data Group** tab, then select the RDG you created.
2. Click **Properties**, then select the **AA Variants** tab.
3. Click **Import**, then navigate to the variant data file. It can be a tab-delimited text file (.txt file).
4. Click **Import**.



5. Select any layer from the drop-down list, and click **OK**.
6. Click **OK** in the Import Results dialog box.


Import Results
6 Variants IMPORTED
0 were reconciled with the reference sequence
0 Variants were NOT IMPORTED
0 not valid
0 duplicated existing variant(s)
There were 0 parsing errors
See the log file for more details.
OK

The amino acid variants are imported and appear in the list in the AA Variants tab. A sample of AA variants is shown below.



Assigning Styles to Variants

The Variant Styles tab allows you to define text coloring styles that identify different types of variants and change the display characteristics of variants in the Project view. Use the Variant Style tab to assign styles to the variants as desired.

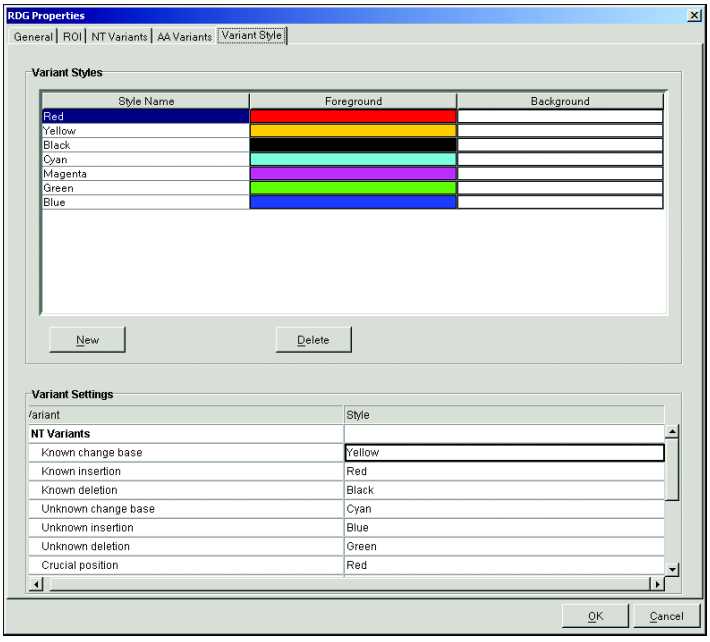
The table at the top of the dialog box displays the generic styles. The table at the bottom of the dialog box lists the different types of variant conditions and their associated styles. The styles you set appear in the Project view to identify the different types of variants.

To assign styles to the variants:

IMPORTANT! When assigning color to text, select light background colors so the text is easy to read.

1. In the RDG Properties dialog box, select the **Variant Style** tab.

   The Variant Styles pane shows the available default colors of the variants.



2. Select the colors in which you want the base changes, insertions, and deletions for known variants displayed.

   a. To add a new color and style, click **New**.

   b. To name the variant style, click the **Foreground Color** box, select a new color in the color palette, then click **OK**.

   c. Select a color from the color palette for the **Background Color**, then click **OK** in both dialog boxes to set the new variant style.

   The variant styles you set appear in the Project view to identify the different types of variants.

3. To delete a color, select the color, then click **Delete**.

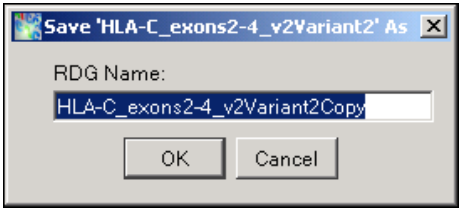
   Note: The first seven Foreground colors cannot be changed or deleted.

4. In the Variant Settings pane, select the colors in which you want the base changes, insertions, and deletions for unknown variants to be displayed. The Variant Styles area shows a list of the available default colors.

Saving a Copy of the RDG

To save a copy of the RDG:

1. In SeqScape Manager, select the RDG you want to save.
2. Click **Save As**.
3. When the confirmation window opens, rename the RDG or click **OK**.



Saving the RDG for Other Projects

If you are working with an RDG that is embedded in a project or project template, you can save a copy of the RDG into SeqScape Manager. This is useful if you make edits to an RDG and want to reuse the RDG for other projects.

To save the RDG:

1. With the project open, in the RDG Properties window, click **Save To Manager As**.
2. Enter a name for the RDG, then click **OK** to save a copy of the RDG under a new name.

If you accept the default name, a copy of the original RDG is saved with that name and is available to use with another project.

Save To Manager As Button

The selections on the Analysis menu, RDG Properties, Analysis Defaults and Display Settings have the Save To Manager As button available for all tabs in each dialog box. The purpose of saving these elements of the project is to have them available to use for changes to the project template for that particular project or other projects that are created.

To use the Save To Manager As button:

1. With the project open, select the **Analysis** menu, then select **RDG Properties**, **Analysis Defaults** or **Display Settings**.
2. Select any tab in any of these three dialog boxes, then click **Save to Manager As**.
3. In the Name field, enter a new name, or accept the default nameCopy and click **OK**.

   The saved copy is available to import into another project.

Creating a Project Template 5

This chapter contains:

Workflow for This Chapter 5-2
Creating a Project Template 5-3
Saving Project Components 5-5

Workflow for This Chapter

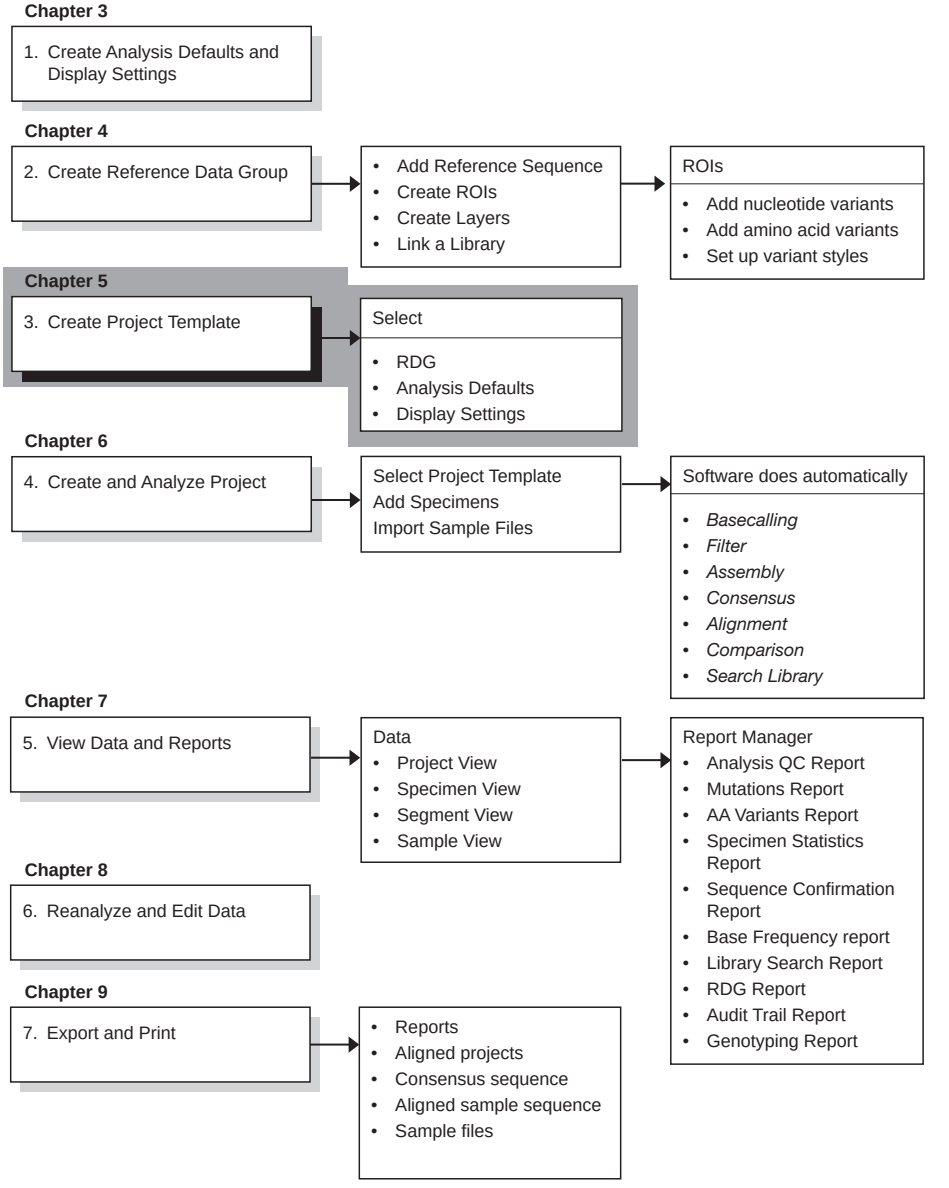


Figure 5-1 Showing Create Project Template Step

Creating a Project Template

Before you can effectively use the ABI PRISM® SeqScape® Software Version 2.1, you must create and configure a project template. A project template contains all the reference data and settings needed to analyze your data automatically. It defines how the software analyzes and displays your samples. When project templates are created in the SeqScape Manager, they can be imported, exported, and edited.

About Creating a New Project Template

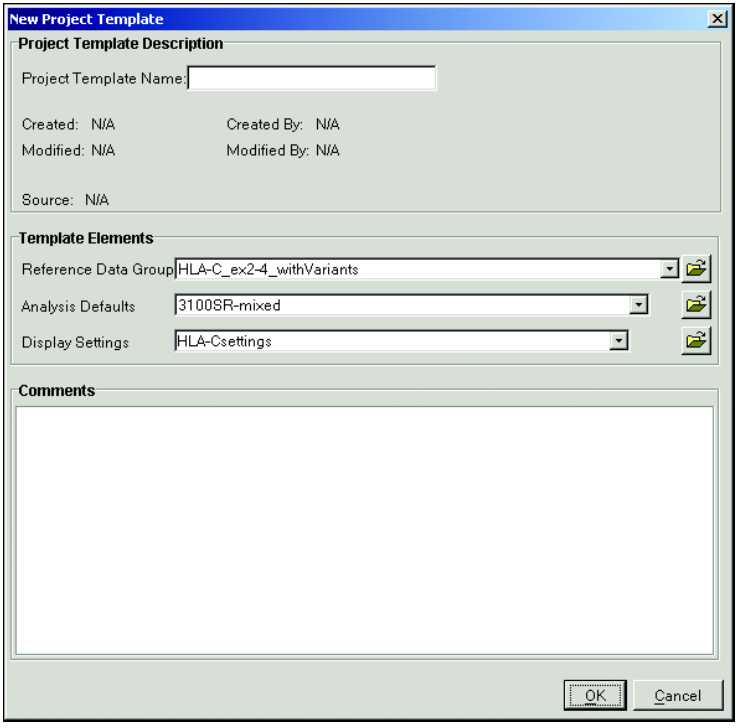
When you create a new project template from the SeqScape Manager, you select:

- A Reference Data Group – Reference sequence and associated data to which all the specimens in a project are compared. See "Creating a New RDG Using SeqScape Manager" on page 4-12 for more information.
- Analysis defaults – Settings that are used to analyze the data. See "Specifying the Analysis Settings" on page 3-12 for more information.
- Display settings – Settings that are used to display the data. See "Specifying Display Settings" on page 3-18 for more information.

Creating a New Project Template

To create a new project template:

1. In the SeqScape window, select **Tools** > **SeqScape Manager**.
2. Select the **Project Templates** tab, then click **New**.



3. Enter a name for the project in the Project Template Name field.

   Note: The project template name must contain only characters that conform to the Windows file system. Refer to "File-Naming Convention" on page 2-10 for a list of all invalid characters.

4. Select the desired Template Elements from the drop-down lists, then click **OK**.

Saving Project Components

About Saving Template Components

If you modify RDG, analysis, or display settings within a project, the changes are valid only in that one project. However, if you want to save those settings so they can be applied to other projects, you can create new SeqScape Manager template components based on existing template components.

Saving Template Components from Within a Project

To save project template components:

1. Within a project, select the Analysis menu, then select one of the template components that you want to modify:
   - **RDG Properties**
   - **Analysis Defaults**
   - **Display Settings**
2. Make the desired modifications to the component.
3. Click **Save To Manager As**.

   An appended name of the current template component appears in the Save.xx to the SeqScape Manager As dialog box.
4. Leave the name unchanged or change it.

   IMPORTANT! You cannot save over an existing template component. You must delete the existing master component from the SeqScape Manager before you can save a new template component.
5. Click **OK**.
6. To use the modified component for other projects, make a new project template that uses the new components.

Examples of Changing the Settings Within a Project

Example 1

1. Create a project template and apply it to a project.
2. Select **Analysis** > **Analysis Settings**, then change a sample analysis setting.

   The underlying Analysis Defaults are unchanged in the SeqScape Manager.

Example 2

1. Create a project template and apply it to a project.
2. Modify each component of the template.
3. Change a variant style in the RDG, then select **Save To Manager As**.

   A new RDG in the SeqScape Manager reflects this change, but the old RDG in the SeqScape Manager remains unchanged. Therefore, the project template using the old RDG is also unchanged.

Note: In both examples, the open project displays the changes.

Creating and Analyzing a Project

6

This chapter contains:

Workflow for This Chapter 6-2
Ways to Create and Analyze a New Project 6-4
Using the New Project Wizard to Create and Analyze a Project . .6-5
Creating and Analyzing a New Project Using a Project Template6-10
Adding Specimens and Importing Data into a Project 6-11
Analyzing the Data 6-21
Reanalyzing a Project Using a Different Project Template. 6-22
Incorporating Variants into the Project RDG 6-25
Importing and Exporting Project Information. 6-34

Workflow for This Chapter

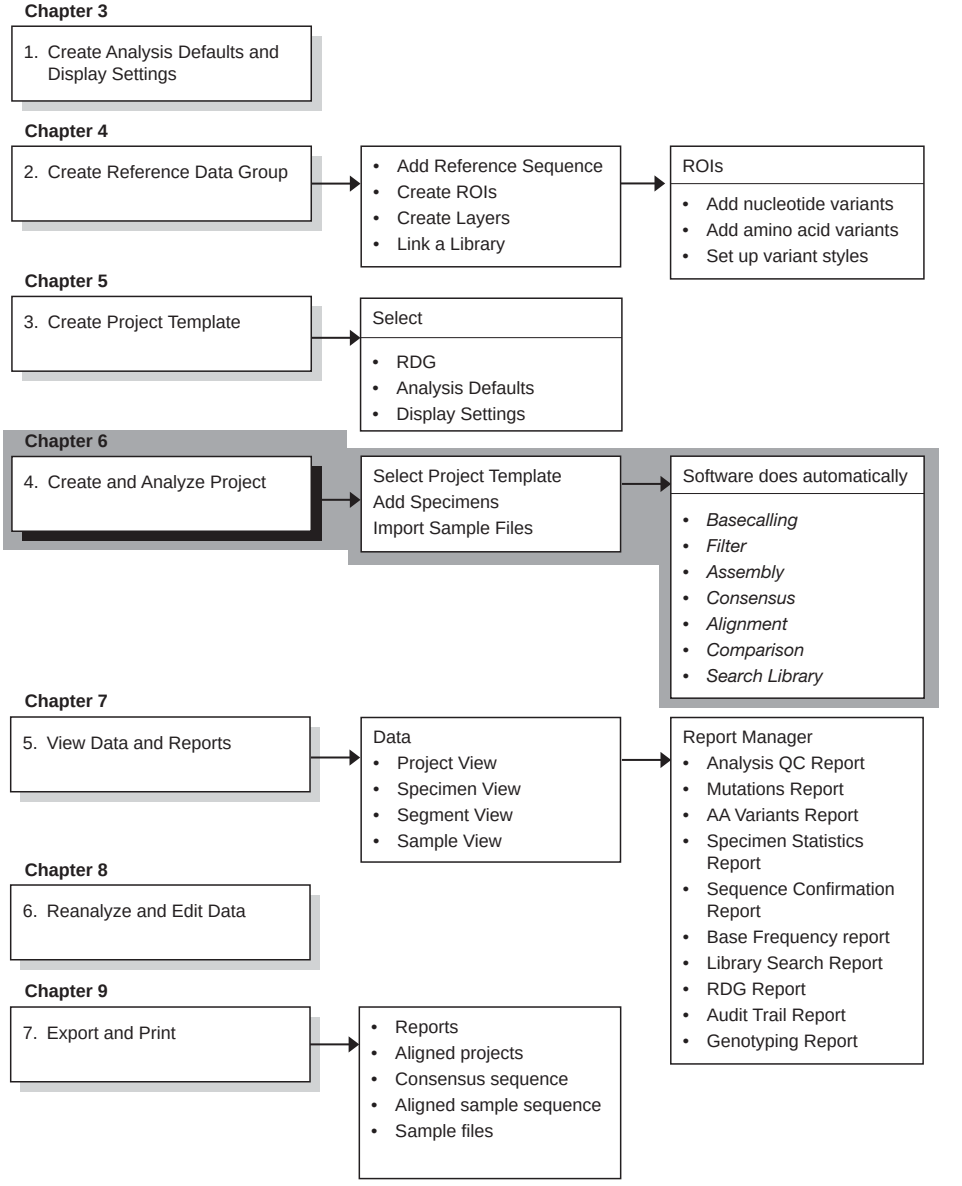


Figure 6-1 Showing Create and Analyze Project Steps

Before You Begin Creating a Project

Before you can create a project in the ABI PRISM® SeqScape® Software Version 2.1, you must have created a project template that contains:

- A Reference Data Group
- Analysis Defaults
- Display Settings

What an Analysis Entails

The analysis in SeqScape software:

- Basecalls the raw data
- Assigns quality values and identifies mixed bases
- Filters out poor quality data and excludes that data from further analysis
- Assembles samples within each specimen to generate a consensus sequence
- Aligns each specimen consensus sequence to the reference sequence
- Compares the aligned consensus sequence to the reference sequence
- Displays analysis results

Ways to Create and Analyze a New Project

After the analysis defaults are set up, you can create a new project for data analysis by:

- Using the New Project Wizard (see page 6-5)
- Using an existing project template (see page 6-10)

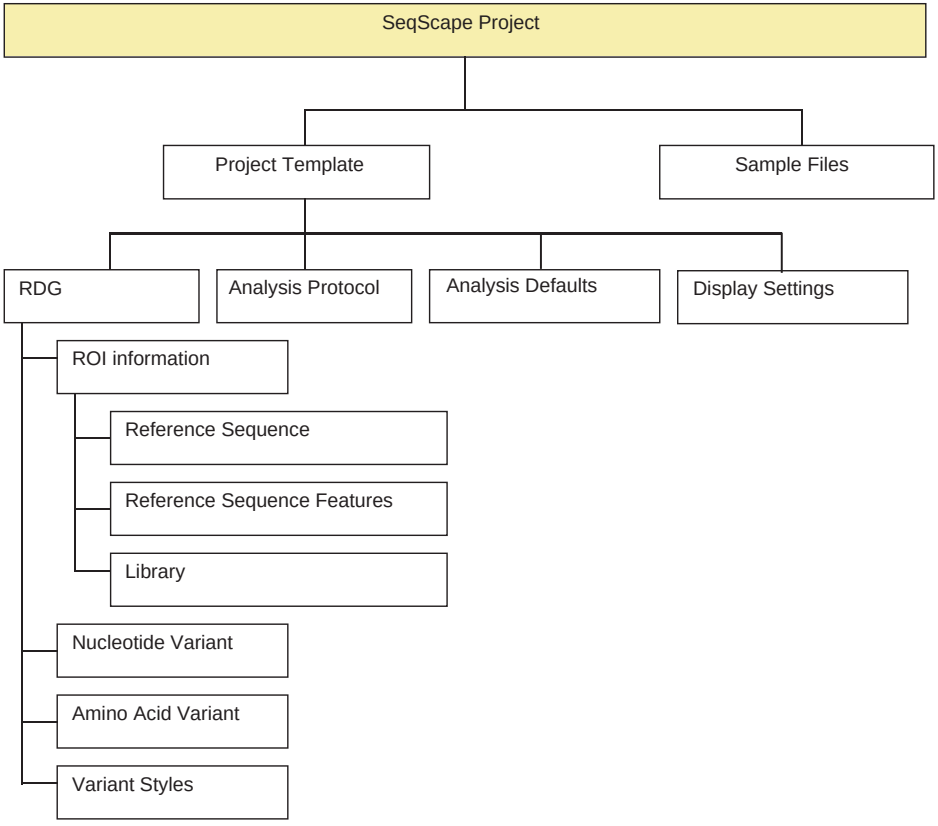


Figure 6-2 Components of a Project

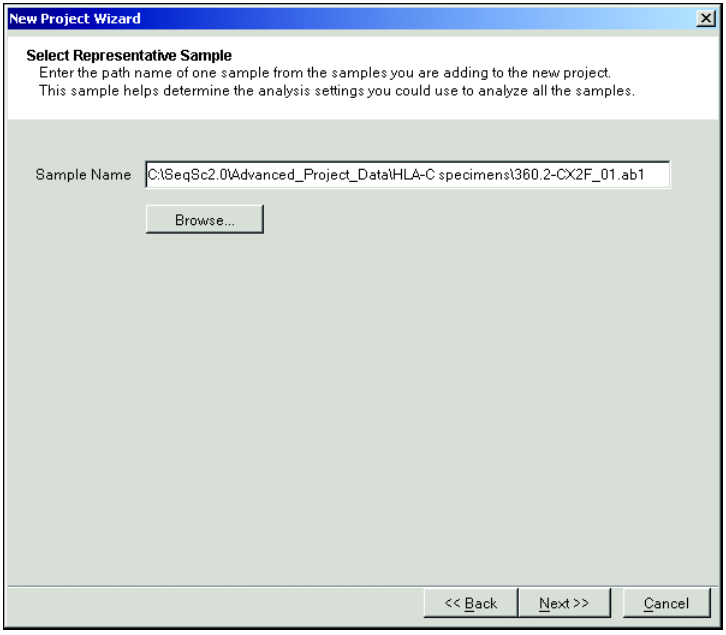
Using the New Project Wizard to Create and Analyze a Project

The New Project Wizard

The New Project Wizard takes you through the process of setting up a new project.

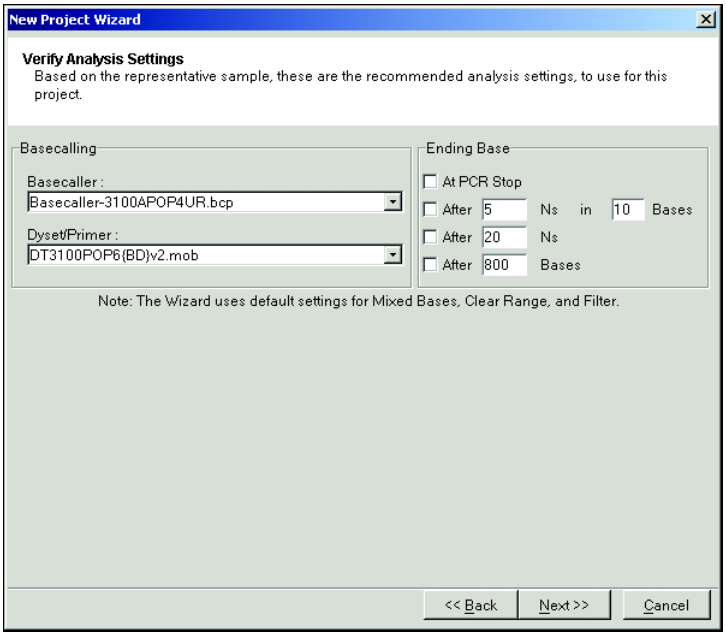
To create a new project using the New Project wizard:

1. Launch SeqScape software.
2. Select **File** > **New Project Wizard**.
3. Enter a name for the new project in the Project Name field, then click **Next**.
4. Enter a sample name, or click **Browse** and navigate to the sample you want.


New Project Wizard
Select Representative Sample
Enter the path name of one sample from the samples you are adding to the new project.
This sample helps determine the analysis settings you could use to analyze all the samples.
Sample Name
C:\SeqSc2.0\Advanced_Project_Data\HLA-C specimens\360.2-CX2F_01.ab1
Browse...
<< Back
Next >>
Cancel

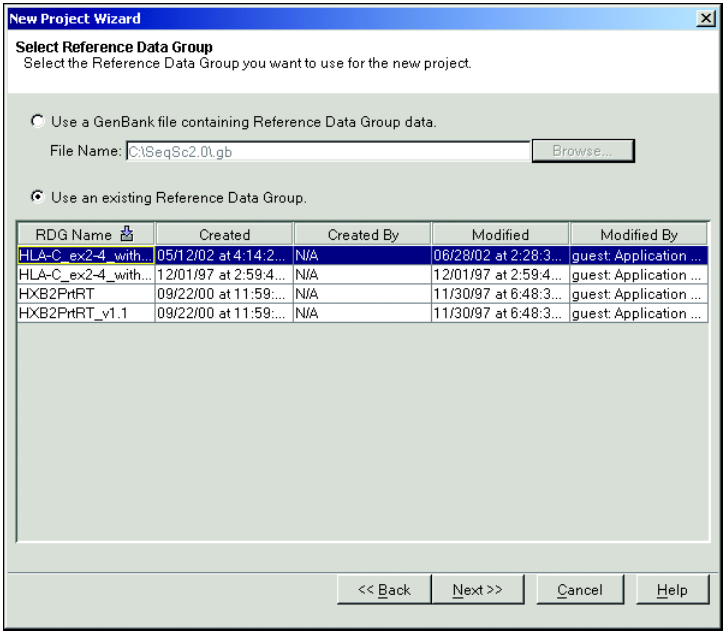
5. Select a sample with the .ab1 extension, then click **Open**.

6. Click **Next**. The wizard uses analysis settings based on your sample choice.



7. Verify the analysis settings (Basecaller, DyeSet/Primer files, and Ending Bases), then click **Next**.

8. In the Select Reference Data Group page:

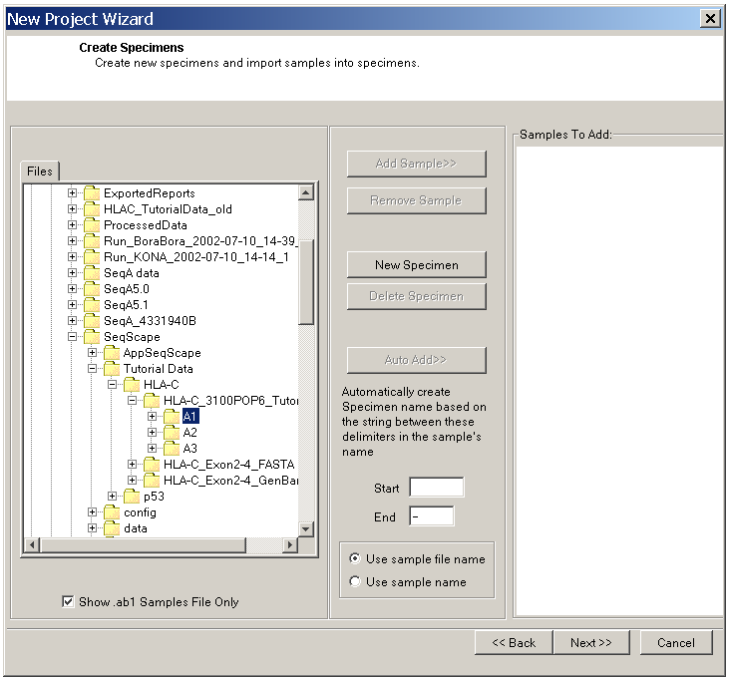


a. Do one of the following:

- Select **Use a GenBank file containing Reference Data Group data**, then specify a GenBank file. Or,
- Select **Use an existing Reference Data Group**, then select a Reference Data Group file in the list.

b. Click **Next**.

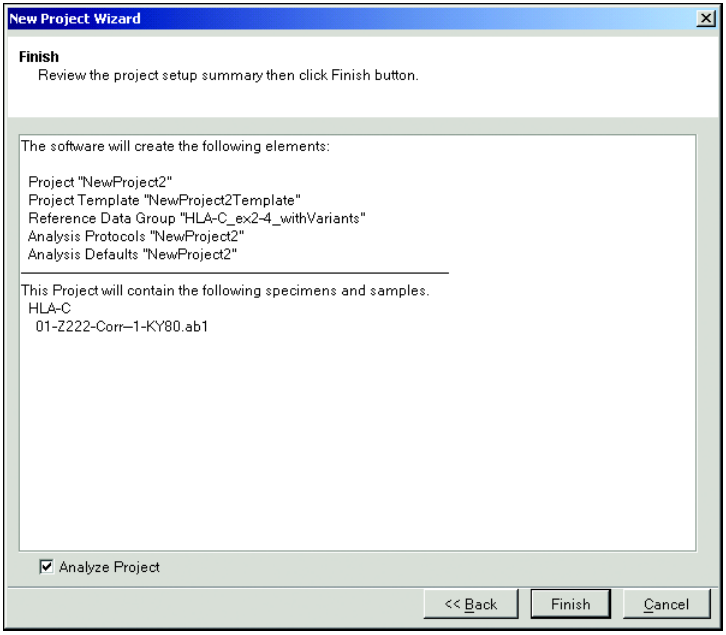
9. Add specimens and import samples in the Create Specimens page:



a. In the Files section, select a sample, multiple samples, or a folder, then click **Auto Add**.

b. Click **Next**.

Note: For information on adding specimens, see "Adding Specimens and Importing Data into a Project" on page 6-11.



When you use the Project Wizard for the first time, master display settings are created. These same settings are used if the wizard is used again.

10. Review the setup. Click **Back** to change the setup, if necessary.
11. Do one of the following:

To analyze …	Then …
Now	Click **Finish**.
Later	1. Deselect **Analyze Project** at the bottom left corner of the page. 2. Click **Finish**.

12. When you close the new project, click **Yes** to save it.

This project is now available in the list of available projects in the SeqScape Manager.

Creating and Analyzing a New Project Using a Project Template

You can use an existing project template to create a new project. A project template contains:

- A Reference Data Group (RDG)
- Analysis defaults
- Display settings

About the Project Template

For convenience, one example project template is included in the software. To create your own project template, see "Creating a Project Template" on page 5-3.

Table 6-1 Components of the Project Template Included in SeqScape Software:

Template Component	File Name
Project Template Name	• HLA-3100_POP6mixed_v2 • p53_exon7-v2
Reference Data Group	• HLA-C_exons2-4_noNT_v2 • p53_Exon7_v2
Analysis Defaults	• 3100-SR_POP6_BDTv1_mixed_v2 • 3700LR_POP5_BDTv1_v2
Display Settings	• DefaultDisplaySettings_v2

Creating a New Project Using a Template

To create a new project using a project template:

1. Select **File** > **New Project**.
2. When the New Project window opens, select a template from the list and enter a project name.

   Note: To see the whole name in the list, click-drag the Project Template heading to the right when the double-headed arrow cursor appears on the column bar.

3. Click **New**.

   The new project using the selected template opens.

Adding Specimens and Importing Data into a Project

Overview

All sample data from a single biological source should be placed inside a specimen within a project. All sample data inside a specimen is assembled, and a consensus sequence is produced. You can think of each specimen as holding the assembled samples from one PCR product, for example. The consensus that is generated is compared to the references and aligned to the other consensus sequences from other specimens.

If you have new, unanalyzed data, you need to create specimens in the project to hold the data. You can add specimens to a project automatically or manually.

For more information on what types of data can be imported into a project, see "Adding Specimens and Importing Samples Manually" on page 6-14.

IMPORTANT! Unanalyzed specimen and sample data show a red slash line through their icons, indicating that analysis is needed.

IMPORTANT! Specimen names can be edited only after they are imported. Sample names cannot be edited from within SeqScape software at any time.

Adding Specimens and Importing Samples Automatically

Using a text delimiter, SeqScape software simultaneously and automatically creates specimens and imports unanalyzed or analyzed samples into a project.

Sample IDs and Sample File Names

To take advantage of this feature, your sample ID (which is created by the data collection software and stored within each sample file) needs to contain the same prefix for all samples in each specimen.

IMPORTANT! The sample ID (sample name) is the name that you assign to the sample in the data collection software. You *cannot* modify the sample name.

The sample file name is longer than the sample ID, and often is derived from the sample ID. The sample file name is what you see when looking for the sample. You can modify the sample file name.

The text delimiter is chosen from the sample ID. Using this function, a set of sample files that are grouped into the same folder and that share a similar delimiter can be imported into their corresponding specimens in a single step.

In the example shown in Figure 6-3, the delimiter is a dash. Everything to the left of the delimiter determines the specimen name. When you select Add Automatically, the sample files are automatically transferred into specimens that are also created and named automatically. In this example, the first specimen includes all files that start with A1.

The sample ID also appears in the Annotation view of the sample, as indicated in Figure 6-3.

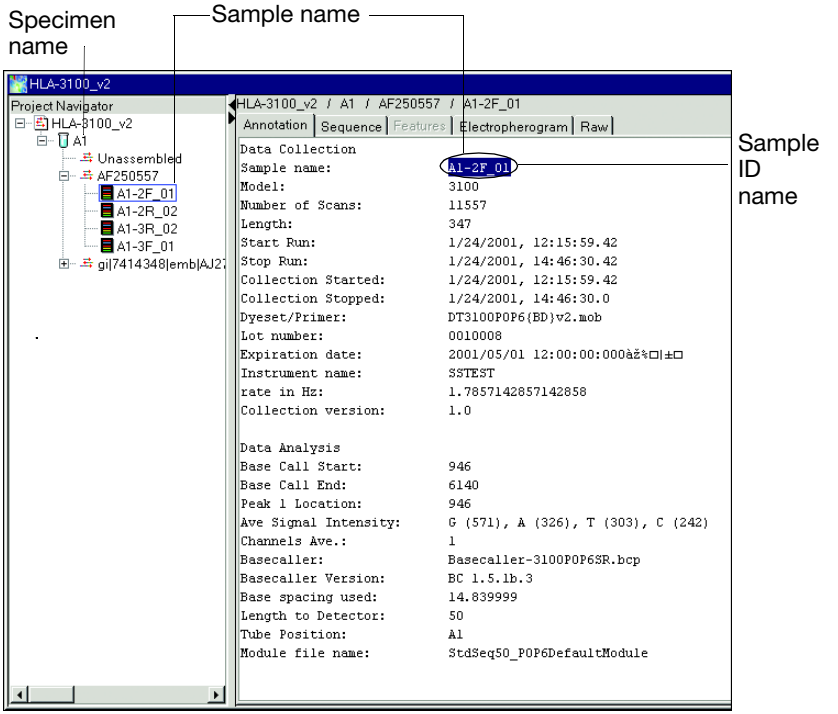


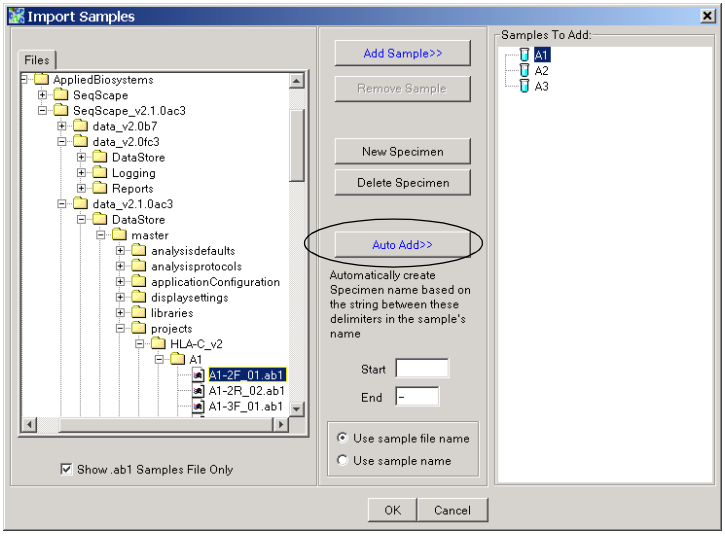
Figure 6-3 Annotation Tab

To create a specimen and import samples automatically:

1. With the Project window open, select **File** > **Import Samples To Project** or click .
2. In the Specimen name delimiter field, enter the delimiter text.

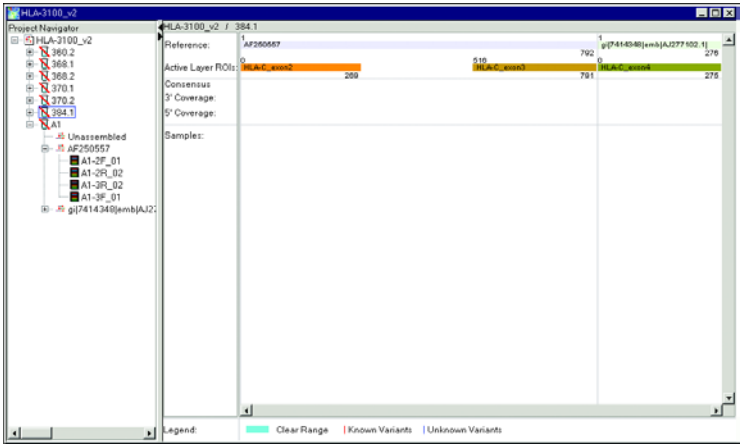
   Note: The delimiter text is derived from the sample ID name in the data collection software sample sheet or plate record. In the figure in step 3, the delimiter is a dash. The sample ID name from the data collection software appears in the Sample name section of the Annotation view of the sample.

3. Select the folder containing the samples to be imported, then click **Auto Add**.



Based on the text delimiter, the samples are automatically imported into the appropriate specimens (in this example, the specimens are shown under HLA-C specimens).

4. Click **OK** to import the specimens and samples into the project.



Adding Specimens and Importing Samples Manually

You can import the following types of sample data into specimens within a project:

- Sample data files from ABI PRISM instruments
- Database files
- Specimen text-only files

Table 6-2 Types of Sample Data

To import …	See …
Sample data files	“Adding Specimens and Importing Data Files” on page 6-15
Specimen text-only files	“Importing Text-Only Files” on page 6-20

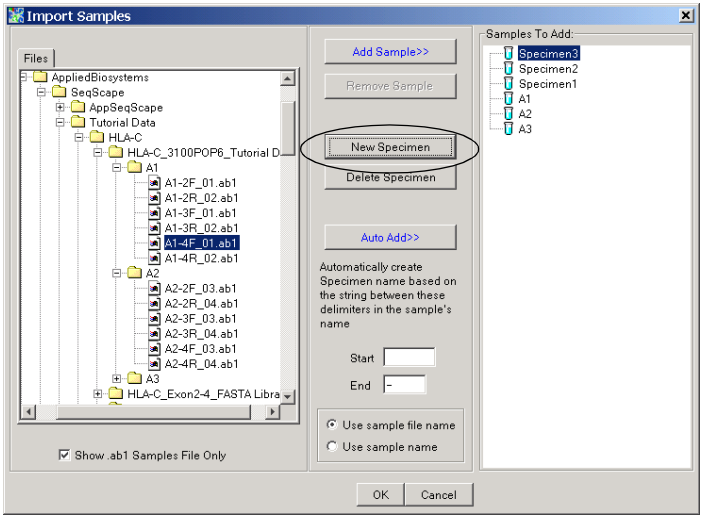
Adding Specimens and Importing Data Files

To import unanalyzed or analyzed sample data, the files must be in ABI format. Sample data is imported into specimens in the project. New specimens are created in the Import Samples dialog box.

To add specimens and import sample data files:

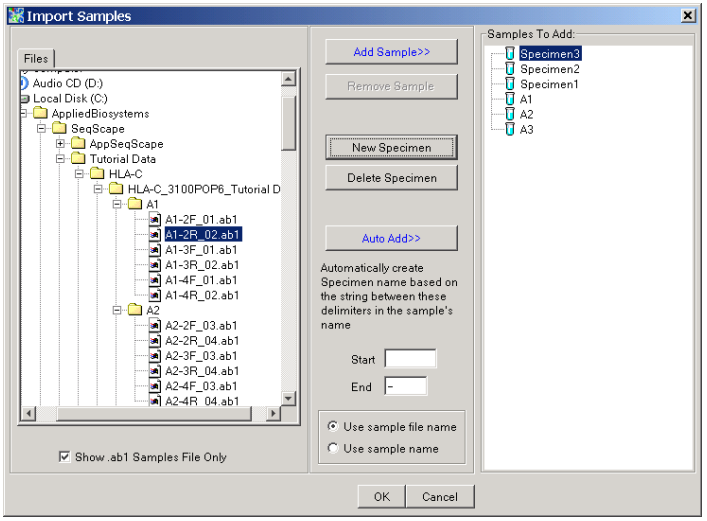
1. With the Project window open, select **File > Import Samples To Project**, or click to open the Import Samples dialog box.

2. Create a new specimen:
   a. Click **New Specimen**.
   b. Add two more specimens.



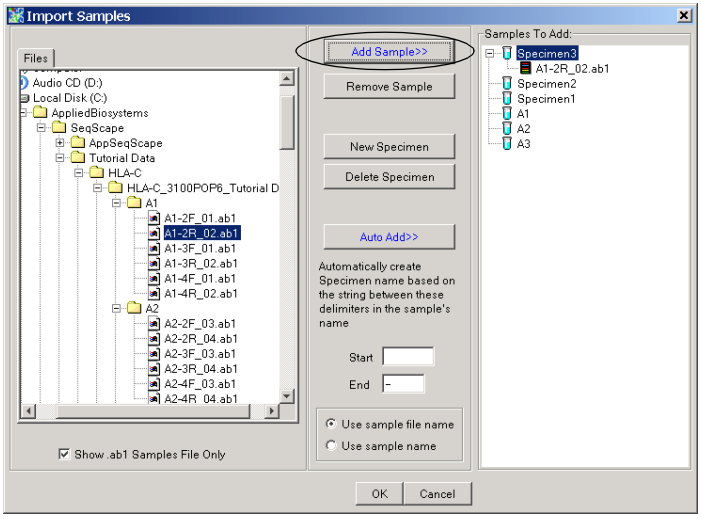
3. In the Samples To Add section on the right, select the specimen into which to import the data.

4. In the Files pane, navigate to the samples you want to add.
5. Select the first specimen in the Samples to Add pane.



6. In the Files section, select the sample data files.

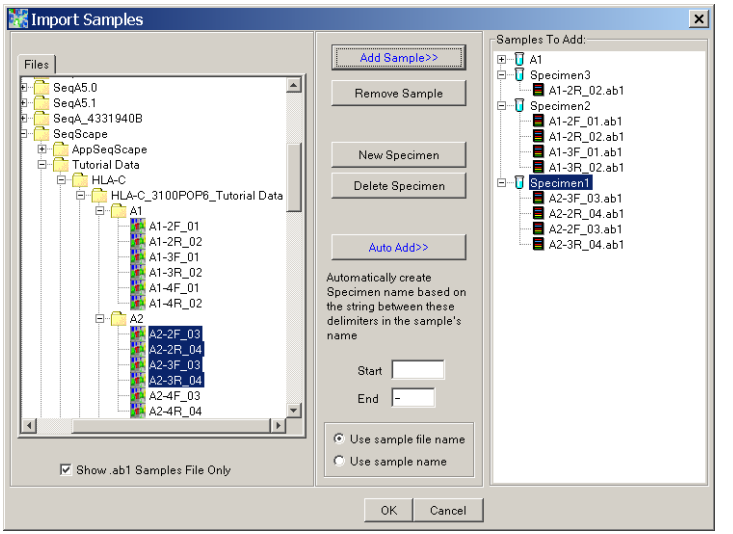
To import …	**Then …**
A single sample	Select the single sample.
Multiple samples	**Ctrl+Click** to select contiguous or noncontiguous samples.
All samples in a folder	Select the folder.



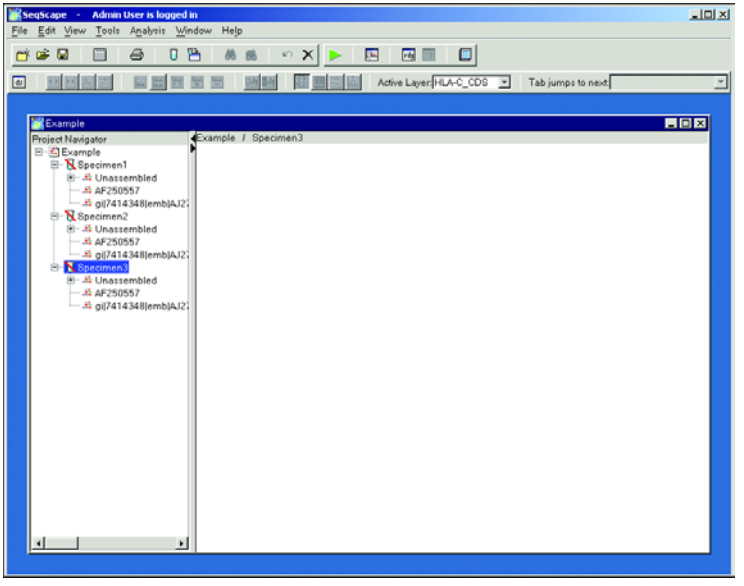
7. Click **Add Sample**.

   The sample data appears in the selected specimen, showing where the data will be imported. No data is actually imported into the project until you click OK.

8. Select the second specimen, select the samples, and click **Add Sample**. Repeat this for the third specimen.

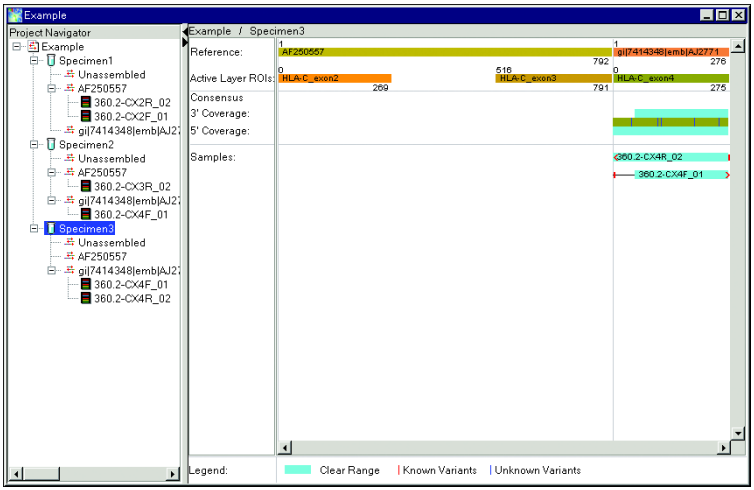


9. Click **OK** to perform the imports and return to the Project window. The project reflects the new specimens and samples, with the specimens shown with a red line through them. This indicates that the samples are unanalyzed and unassembled.



10. If desired, select each specimen and type a new name for the specimen, then press the **Enter** key.

11. The green arrow button on the toolbar indicates that the samples need to be analyzed. Click this button. After the samples are analyzed, the red line through the specimen is gone and the samples are assembled as shown in the figure below.



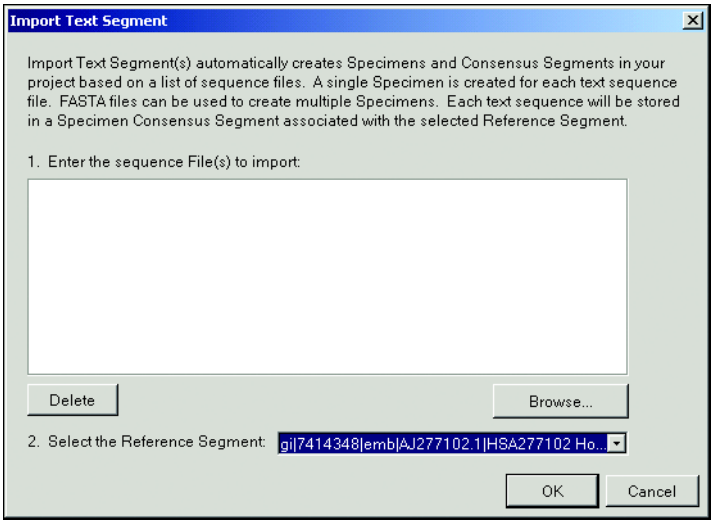
12. Close the project and save it.

Importing Text-Only Files

You can import into a project a consensus sequence in text format as a text-only specimen.

To import text or previously assembled sequences:

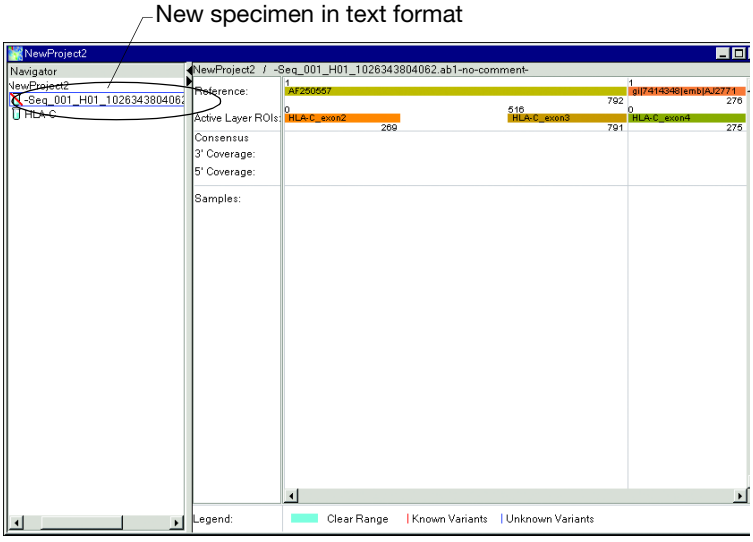
1. In the Project Navigator, select the project name, then select **File > Import Text Segment**.


Import Text Segment
Import Text Segment(s) automatically creates Specimens and Consensus Segments in your project based on a list of sequence files. A single Specimen is created for each text sequence file. FASTA files can be used to create multiple Specimens. Each text sequence will be stored in a Specimen Consensus Segment associated with the selected Reference Segment.
1. Enter the sequence File(s) to import:
Delete
Browse...
2. Select the Reference Segment:
gi|7414348|emb|AJ277102.1|HSA277102 Ho...
OK
Cancel

2. Click **Browse** and navigate to the target segment.
3. In the Import Text-Only Segment dialog box, select the text file (.fsta format), then click **Import**. The segment appears in the previously blank section in the Import Text Segment dialog box.
4. Repeat steps 2 and 3 to add additional text segments.

5. Click **OK**.

   A new specimen is created with the name specified in the first line of the file.



Removing Samples or Specimens

To remove samples or a specimen from a project:

1. In the Project Navigator of the project, select the samples or specimen you want to remove from the project.
2. Press the **Delete** key.

   IMPORTANT! This deletes the results and cannot be undone. If you press Delete in error, close the project without saving to restore the results.

3. In the Confirm Delete dialog box, click **Yes**.

Analyzing the Data

After you import all your data, you can run the analysis. After new data is imported or analysis settings are changed for a sample, the Analyze icon in the toolbar appears green, indicating that there is unanalyzed data.

Running an Analysis

To run an analysis in the project, click (Analyze), or select **Analysis** > **Analyze**.

Reanalyzing a Project Using a Different Project Template

When You Would Want to Do This

After you analyze an entire project that contains many samples, you may want to reanalyze all the data using a project template that contains different settings or reference data.

Saving a Project Before Reanalyzing

IMPORTANT! Applying a new project template discards all analyzed data, including basecalls, features, alignments, and manual edits. To avoid discarding the data, rename the project to keep your original analysis, if desired.

To save a project that you want to reanalyze:

1. Select **Tools** > **SeqScape Manager**.
2. In the Project list, select the project that you want to save before reanalyzing.
3. Click **Save As** and rename the project.
4. Click **OK**.

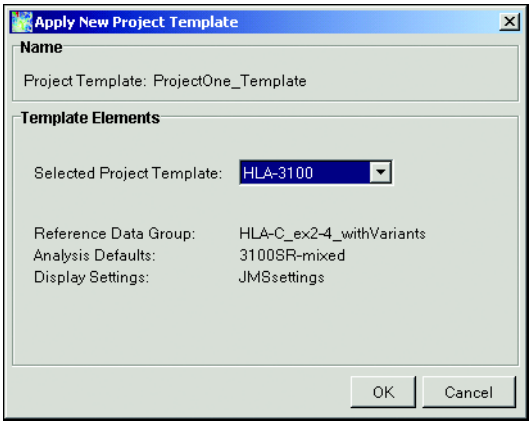
   The project is saved under a new name and your original project remains in the list.

Applying a Template to an Existing Project

To reanalyze the project with a different template:

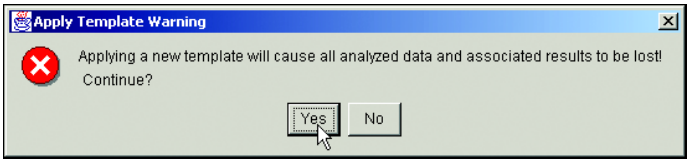
1. Create a template containing the desired changed settings and/or reference sequence (see “Creating a New Project Using a Template” on page 6-10).
2. Open the existing project that has the data analyzed using the old settings.

   IMPORTANT! Save the project under a new name if you want to keep the current project data to compare to the new project data. If you do not save the project, all the data is overwritten when you apply a new project template.
3. Select **Analysis > Apply Project Template** to open the Apply New Project Template dialog box.



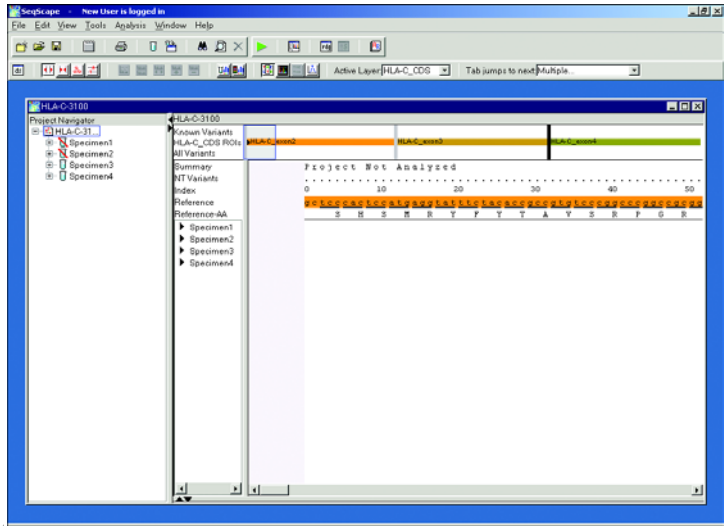
4. In the Selected Project Template drop-down list, select the project template that you want to apply to the project.
5. Click **OK**.

6. A dialog box opens warning you that all analyzed data and results will be discarded. To continue, click **Yes**.


Apply Template Warning
Applying a new template will cause all analyzed data and associated results to be lost!
Continue?
Yes
No

The project opens, containing all the specimens and samples, but the data is unanalyzed.

7. To analyze the data with the new template, select **Analysis** > **Analyze**.



Incorporating Variants into the Project RDG

About Incorporating Variant Sequences

You can incorporate variants into an active project RDG by doing one of the following:

- Changing an unknown variant in a specimen to a known variant
- Adding a variant
- Importing a file containing variant sequences
- Importing a set of variants from a TXT file

Note: If you have a master RDG and want to include additional variants in the RDG, you must incorporate them using the SeqScape Manager.

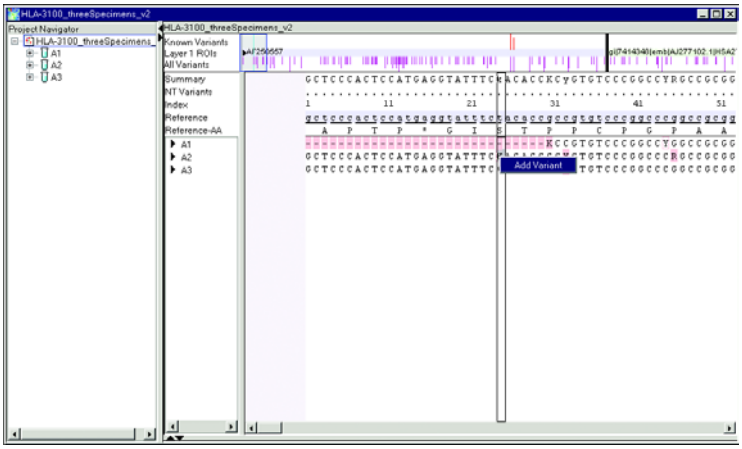
You can incorporate variants automatically by importing a file of a tab-delimited text file of variant positions and descriptions. By this method, variants are created and styles are applied to all the variants in the file.

Alternatively, you can change an unknown variant in a specimen to a known variant, or you can create variants by adding them to the Reference Data Group.

Changing a Single Unknown Variant to a Known Variant

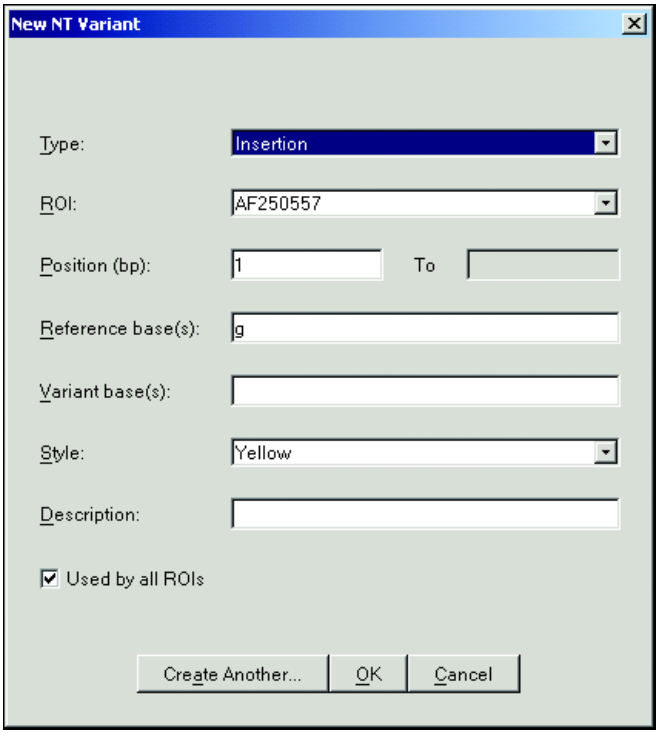
To change an unknown variant to a known variant:

1. Select a variant base in a specimen, then right-click it.



2. Select **Add Variant** from the shortcut menu to open the New NT Variant dialog box.

   The information regarding the type and position of the variant appears in the New NT Variant dialog box.



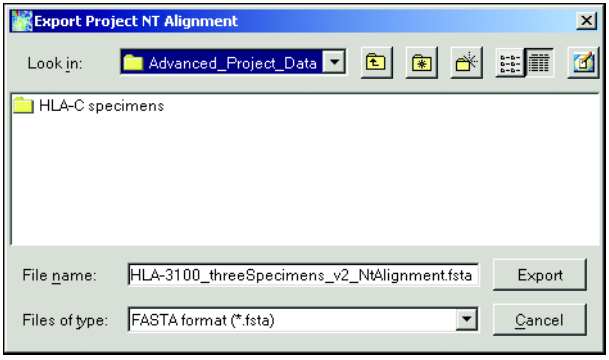
3. Select a variant style from the Style drop-down list.
4. In the Description field, enter text, if desired, then click **OK**.
5. Repeat steps 1 through 4 for another variant.

Changing Multiple Unknown Variants

To change multiple unknown variants to known variants, you need to export unknown variants in a project alignment file and then import them into the project.

To change multiple unknown variants to known variants:

1. Open the project and select the specimen containing the unknown variants.
2. Select **File > Export > Project Alignment-Nucleotides**.

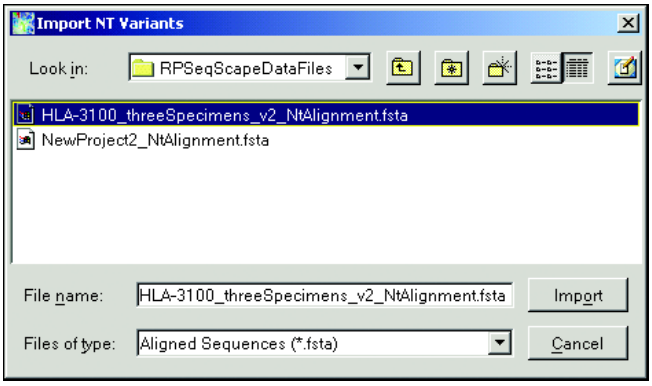


3. In the Export Project NT Alignment dialog box, select a destination for the exported data, then click **Export**.

Importing Variants

To import variants:

1. Select **Analysis > RDG Properties**.
2. Select the **NT Variants** tab.
3. Select **Import**.
4. In the **Import NT Variants** dialog box, navigate to the project alignment file, then select it. Make sure the Files of type is set to **All Files** or **Aligned Sequences**.



5. Navigate to or select the file to import.
6. Click **Import**.
7. Select the reference segment in the drop-down list in the Select Reference Segment dialog box, then click **OK**.

   The variants appear in the NT Variants table as Known variants. The descriptions are the specimens in which the variants appear and the style is the default style for the variant type.
8. Select the **Variant Style** tab in the RDG Properties dialog box to change the default style in the RDG and enter a description of the imported variants.
9. Select the **NT Variants** tab to be sure the variants are Known.
10. Click **OK** to save the variants.

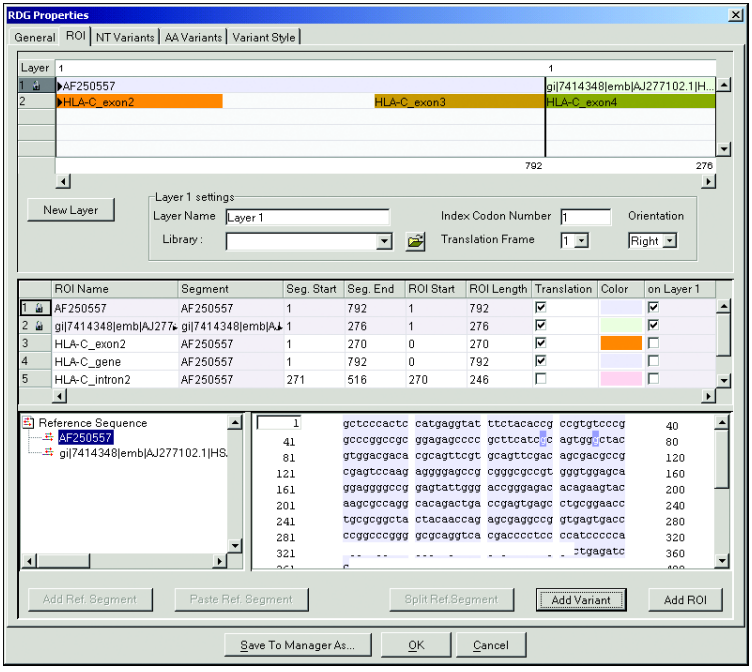
Creating a New Variant in a Project

You can add a variant to a project by:

- Entering the type and position of the variant in the Variants tab of the RDG Properties dialog box.
- Selecting the location on the reference sequence in the Sequence tab in the RDG Properties dialog box. The appropriate information regarding the variant is automatically entered in the variant dialog box.

To create a new variant in the project:

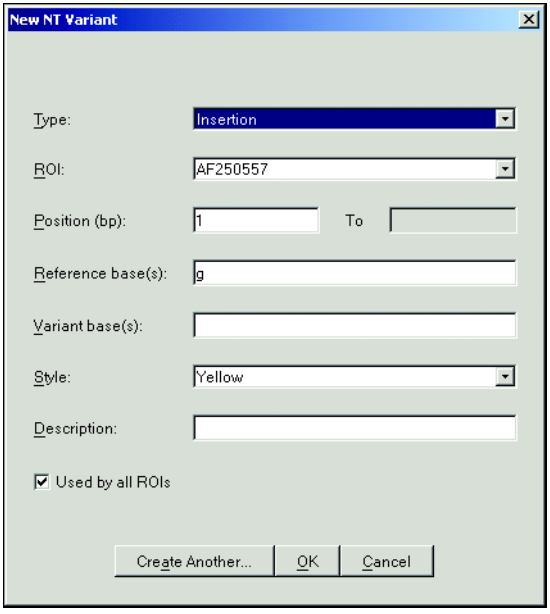
1. From the Project window, open the RDG Properties dialog box by selecting **Analysis > RDG Properties**.
2. Click the **ROI** tab, then select **Add Variant**.



3. In the New NT Variant dialog box, select the type of variant (**Base Change**, **Insertion**, or **Deletion**).

4. Select the Position and either **To** (position) or **Variant base**.

   Note: The **Reference** base is entered by the software based on the position.

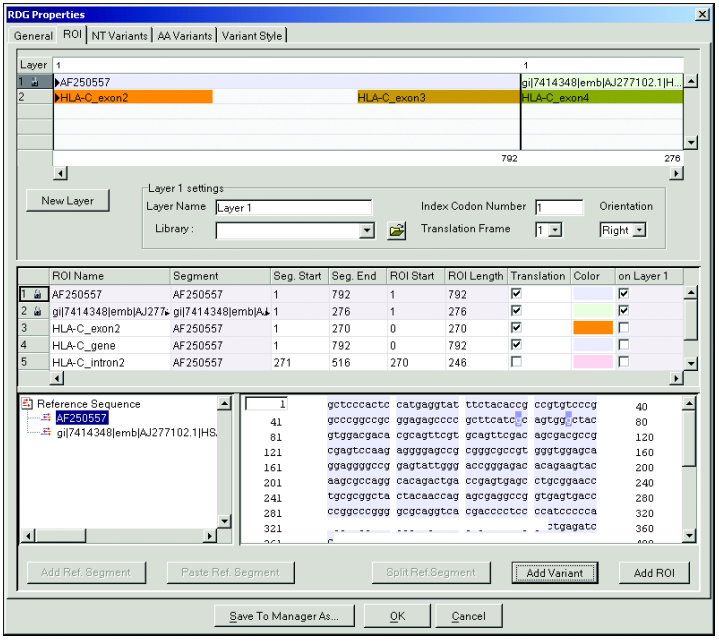


5. Select a style (color) with which you want the variant to be displayed, then enter a description of the variant, if desired.
6. Click **Create Another** to add more variants, or click **OK** to save the variant to the RDG.

Adding a Variant in the Project

To add a variant in the project:

1. From the Project window, select **Analysis** > **RDG Properties**.
2. Select the **ROI** tab.
3. Indicate your variant by doing one of the following:
   - Select the base that corresponds to the substitution variant or range of bases for a deletion variant.
   - Click the position at which you want an insertion variant.
4. Select **Add Variant**.



5. In the New NT Variant dialog box, note that the Position and the Reference base are already entered.
6. Select the type of variant by clicking **Base Change**, **Insertion**, or **Deletion,** then enter the Variant base.
7. Select a style for the variant, then enter a description of the variant, if desired.
8. Click **OK** to save the variant to the project.

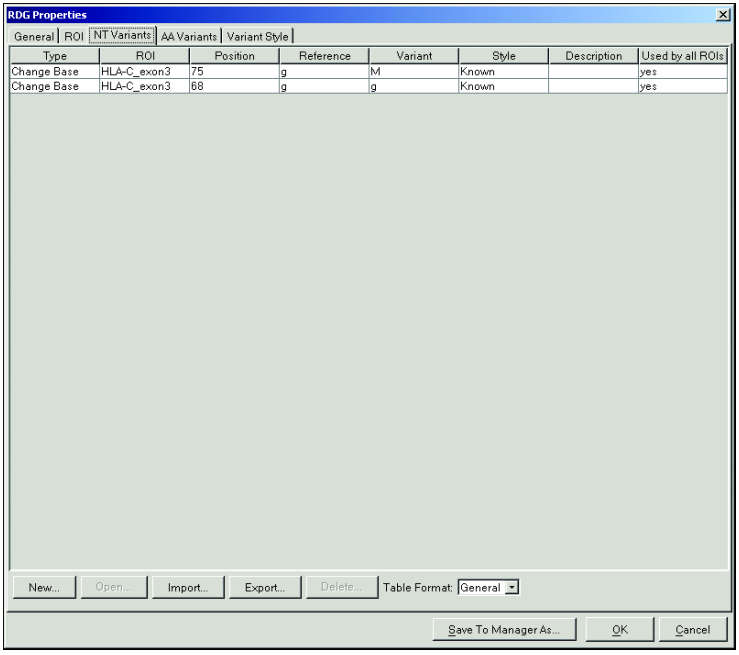
Importing Variants to the Project

When you import variants into a project, they must be in one of the following configurations:

- Tab-delimited text file format
- Text file format containing aligned sequences

To import variants into a project:

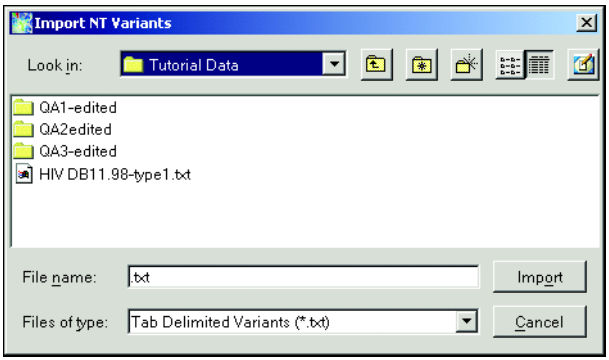
1. Open the RDG Properties dialog box and select the **NT Variants** tab.
2. Click **Import**.



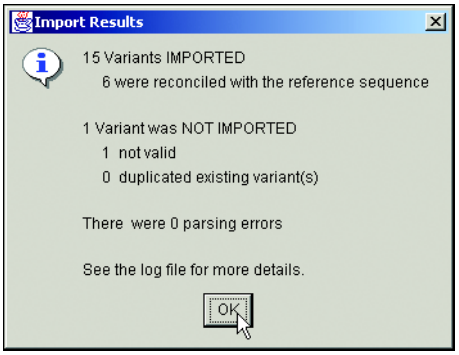
3. Browse to the appropriate file and select it.

   Note: The files must be tab-delimited text files as indicated in the File Types field.

4. Click **Import**.



5. Select the reference segment, then click **OK**.
6. After the data is imported, the Imports Results dialog box opens with information regarding the import.
7. Note the information, then click **OK**.


Import Results
15 Variants IMPORTED
6 were reconciled with the reference sequence
1 Variant was NOT IMPORTED
1 not valid
0 duplicated existing variant(s)
There were 0 parsing errors
See the log file for more details.
OK

The variants now appear in the NT Variants tab of the RDG Properties dialog box.

Importing and Exporting Project Information

About Importing and Exporting

The purpose of importing and exporting project information is to transfer the project information to another computer.

You can export or import projects, project templates, reference data groups, nucleotide and amino acid variant tables, libraries, and analysis defaults from the SeqScape Manager. This allows you to examine and compare results from different Data Stores.

Note: The export and import functions of SeqScape Manager use the file extension CTF.

Importing from SeqScape Manager

To import from SeqScape Manager:

1. Select **Tools** > **SeqScape Manager**.
2. Select any tab into which you want to import.
3. Click **Import**.
4. Navigate to the file that you want to import.
5. Click **Import**.

   The imported file appears in the list under the appropriate tab.

Exporting from SeqScape Manager

To export from SeqScape Manager:

1. Select **Tools** > **SeqScape Manager**.
2. Select any one of the tabs from which you want to export.
3. Select the file that you want to export from the list, then click **Export**.
4. Navigate to a location to export.
5. Rename the file, if necessary, using the .ctf extension.
6. Click **Export**.

   The exported file is available to import into another project.

Viewing the Results 7

This chapter contains:

Workflow for This Chapter 7-2
About the Project Data 7-3
Project Views. 7-4
Specimen Views 7-10
Segment Views 7-11
Sample Views 7-15
Viewing Variants 7-20
About the Reports 7-22
Viewing the Reports 7-23
Viewing the Reports and Project Results 7-25
Analysis QC Report 7-26
Mutations Report. 7-28
AA Variants Report 7-31
Specimen Statistics 7-32
Sequence Confirmation Report 7-33
Base Frequency Report 7-34
Library Search Report 7-35
RDG Report. 7-37
Audit Trail Report 7-38
Genotyping Report 7-39
Customizing the Reports 7-42

Workflow for This Chapter

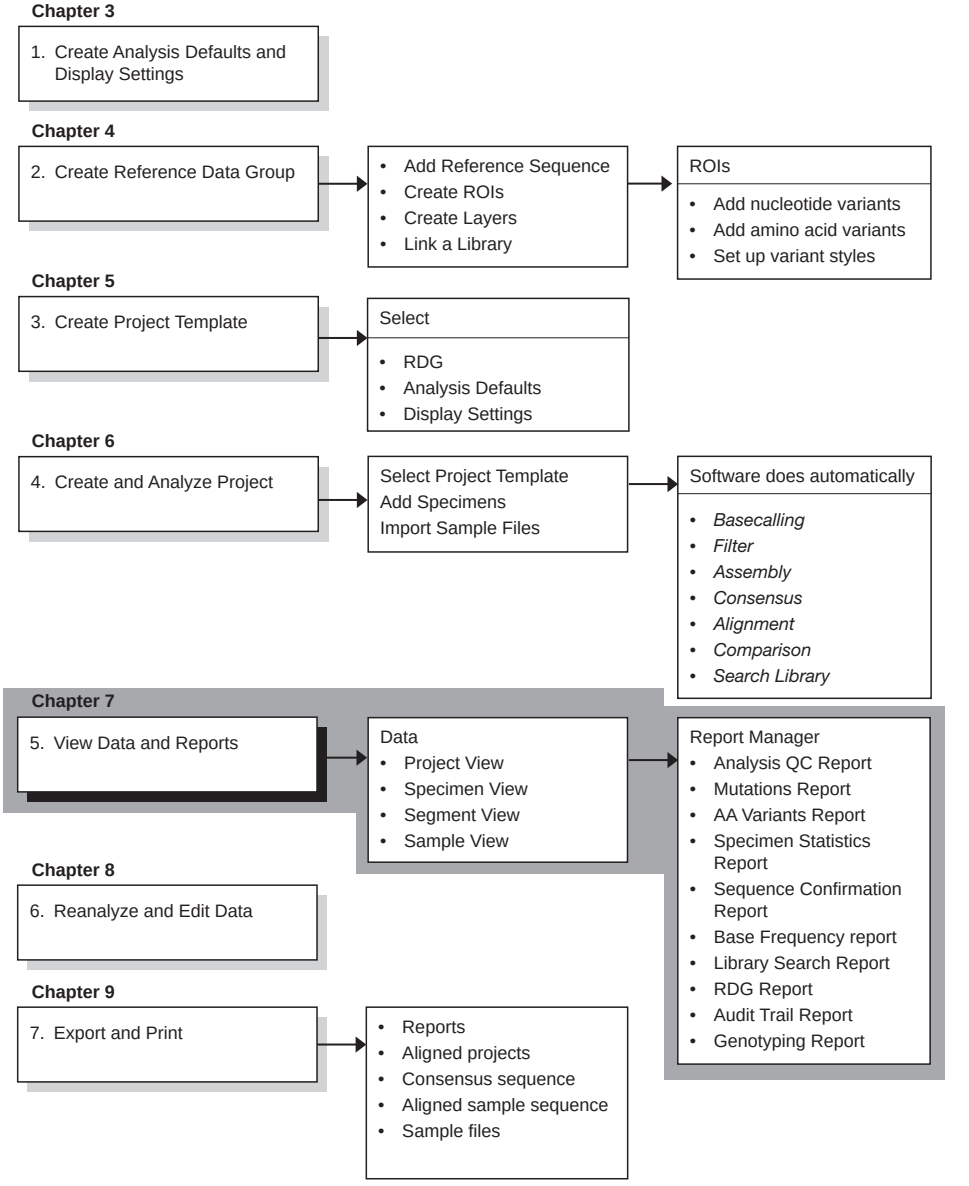


Figure 7-1 Showing the View Data and Reports Steps

About the Project Data

View Formats

You can view the results in multiple formats:

- Project view – A summary of all the specimen consensus sequences
- Specimen view – A summary of all the segment sequences within each specimen
- Segment view – A summary of all the sample sequences within each segment
- Sample view – A summary of the data for each sample

Data Display Conventions

The sequence data is displayed using the following conventions:

- Every mixed base (or choice of mixed bases) is represented as a single IUB code. For more information, see Appendix C, "Translation Tables."
- Spaces in aligned sequences are displayed as dashes and are not part of the original sequence.
- In the Dots view and in the collapsed NT view, characters that are identical to the reference are displayed as dots.
- The aligned reference sequence appears at the top of the table and the aligned sequences appear in the rows below in the Project view.

Quality Value Display

The QV (quality value) is displayed as a bar above each called base for the sample sequence and consensus sequence. The height of a bar corresponds to a 1–99 value that is determined by the analysis.

Note: For QVs from 50 to 99 all QV bars are identical in height and color.

Note: For more information on quality values, see Chapter 10, "Sample and Consensus Quality Values."

Exporting and Printing Project Data

To export the project data, see "Exporting Data Files" on page 9-3, and to print data, see "Printing Data and Reports" on page 9-11.

Project Views

There are three project views, only one of which can be displayed at a time:

- Expanded NT
- Collapsed NT
- Expanded AA

Displaying the Project Views

To display project views:

1. Open the project of interest.
2. Select a layer in the Active Layer drop-down list.
3. At the top of the navigation pane, select the project icon.
4. Use the instructions in Table 7-1 to display the project views of interest.

Table 7-1 Project Views

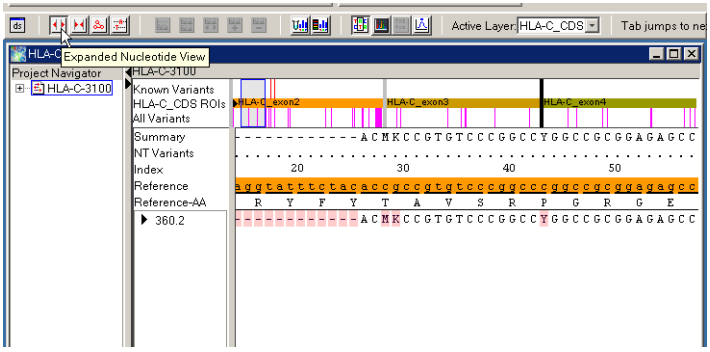
View	Procedure	Display
Expanded Nucleotide	Click .	

Table 7-1 Project Views *(continued)*

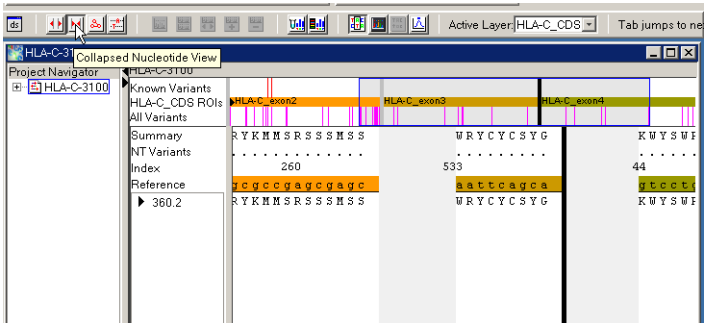
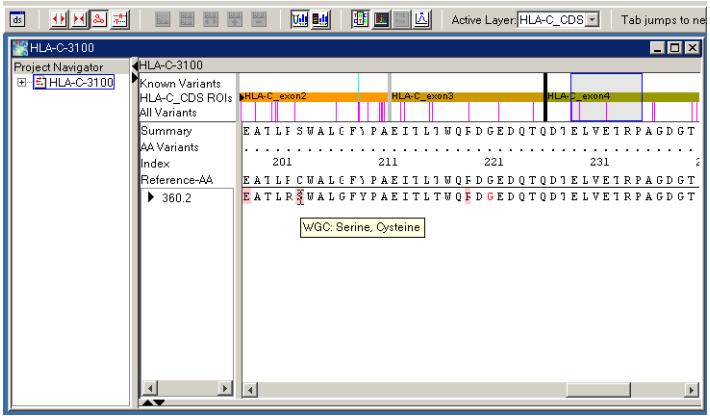
View	Procedure	Display
Collapsed Nucleotide	Click .	 This view shows only those columns that differ from the aligned reference. Bases that match the reference sequence are displayed as dots, regardless of the state of the Dots setting. **Note:** Click (Expanded Nucleotide View) to return to the expanded view.
Expanded Amino Acids (translation of the nucleotide sequence)	Click .	 **Note:** Holding the pointer over an amino acid displays the possible translations and the codon at that position. **Note:** Bold red characters (default) indicate the location of a degenerate codon.

Table 7-1 Project Views *(continued)*

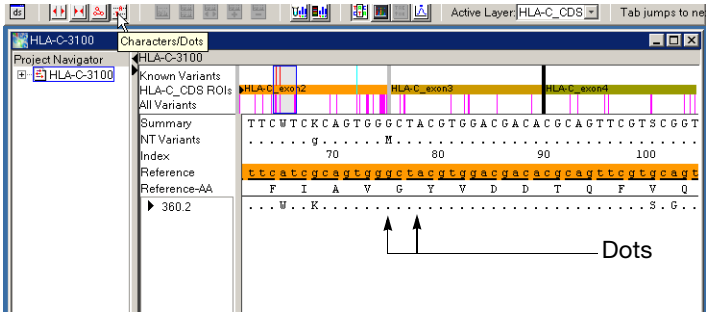
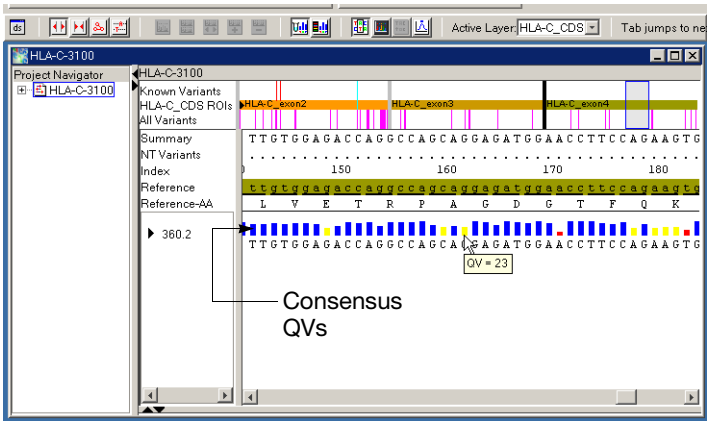
View	Procedure	Display
Dots	Click .	 **Note:** The characters that are identical to the reference are displayed as dots.
Consensus QV	1. Click . 2. Click .	 **Note:** Holding the pointer over a base displays the numerical QV assignment for that base.

Table 7-1 Project Views *(continued)*

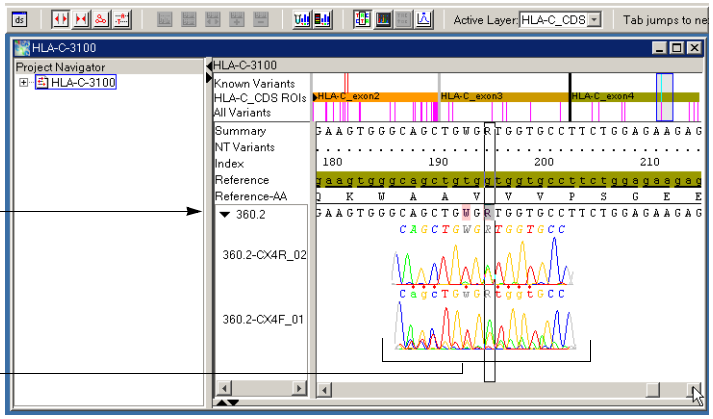
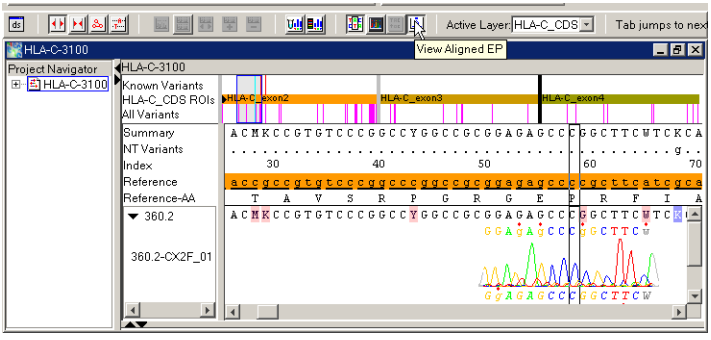
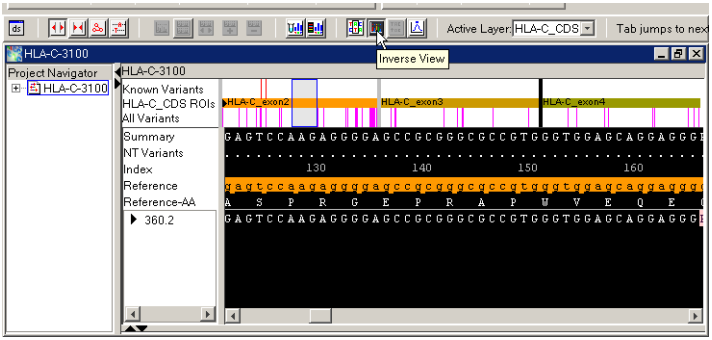
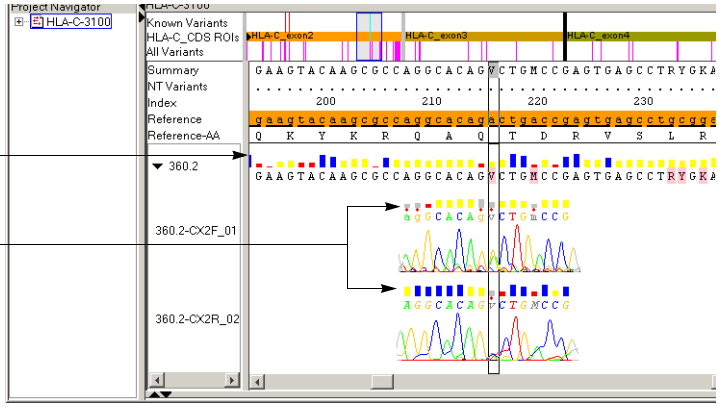
View	Procedure	Display
Electro-pherogram Snippet	1. In the Expanded Nucleotide or Dot view, select a base in the summary or specimen sequence. 2. Click the triangle next to the specimen name. Triangle Electropherogram snippets **Note:** Pressing Ctrl+Z centers any electropherogram snippets in the middle of the view.	
View Aligned EP	Click .	

Table 7-1 Project Views *(continued)*

View	Procedure	Display
Inverse Video	Click .	
QVs with Snippets or Dots	From the Snippet or Dots view, click for sample QVs and/or click for consensus QVs.	

Consensus QVs

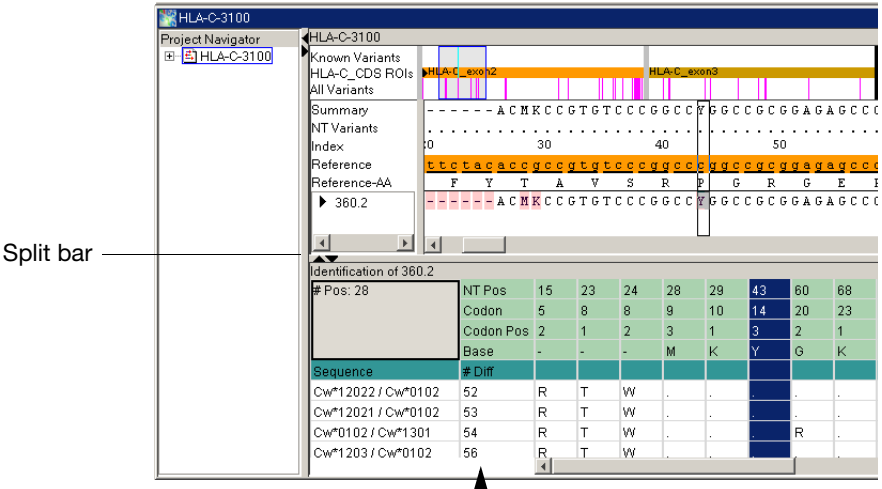
Sample QVs

Note: A gray QV bar indicates that the base has been edited and that the QV no longer holds true (until you restore the previous base).

Note: Holding the pointer over a base displays the numerical QV assignment for that base.

Table 7-1 Project Views *(continued)*

View	Procedure	Display
Identification	1. Drag the split bar at the bottom of the window up until you reach the desired height. 2. Select a base in a specimen sequence.	

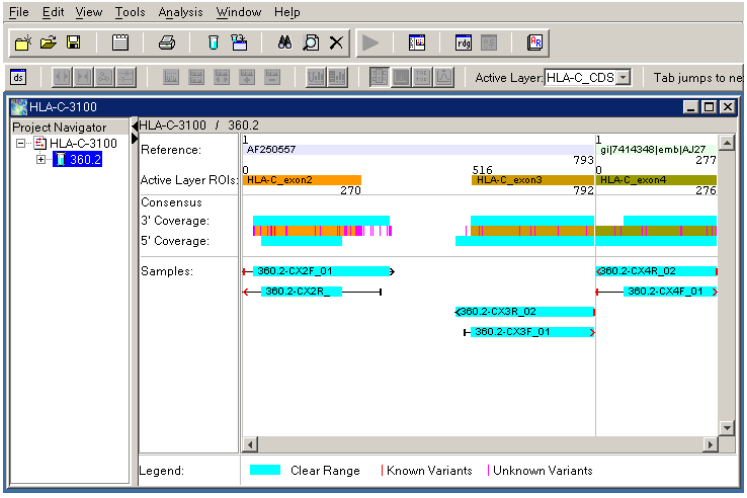


Specimen Views

The specimen result is displayed as a schematic of the location and orientation of all samples within a specimen with respect to the reference, ROIs in the current layers, and consensus sequence.

To display the specimen view:

1. Open the project of interest.
2. Select a layer in the Active Layer drop-down list.
3. In the navigation pane, open the project (if necessary), then select a specimen icon.



Segment Views

There are two segment views:

- Layout view – Displays a schematic of the location and orientation of the samples with respect to the consensus segment and the reference
- Assembly view – Displays the nucleotide sequence of the consensus and samples, sample electropherogram data, and view position.

Note: The view position in the Assembly view (blue box) is represented by red lines in the Layout view. Click the Layout view to navigate to a desired position in the Assembly view.

Table 7-2 describes the multiple Assembly view types.

Displaying the Segment Views

To display segment views:

1. Open the project of interest.
2. Select a layer in the Active Layer drop-down list.
3. In the navigation pane, open a specimen, then select a segment.
4. Use the procedures in Table 7-2 to display the segment views of interest.

Table 7-2 Segment Views

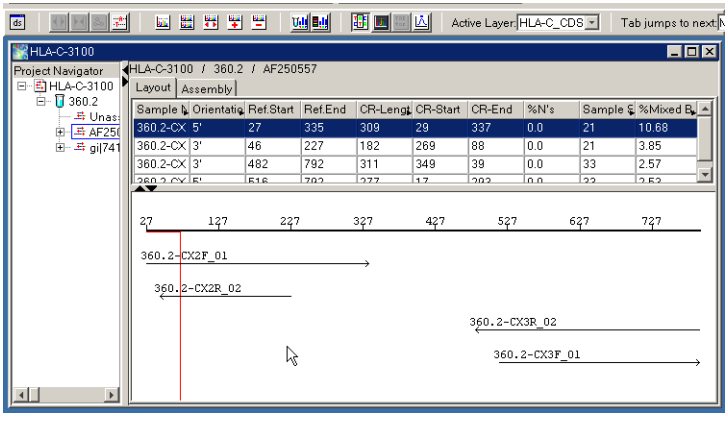
View	Procedure	Display
Layout	Select the **Layout** tab.	

Table 7-2 Segment Views *(continued)*

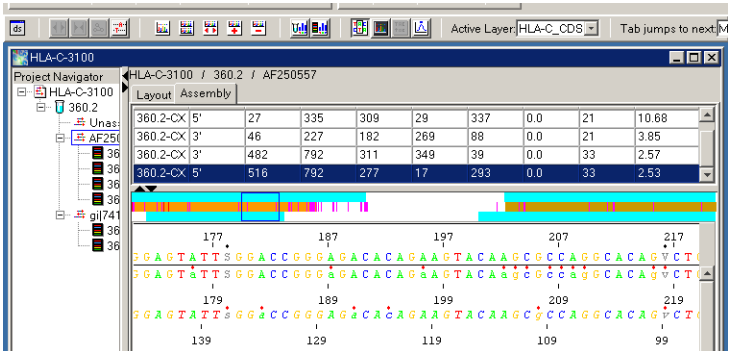
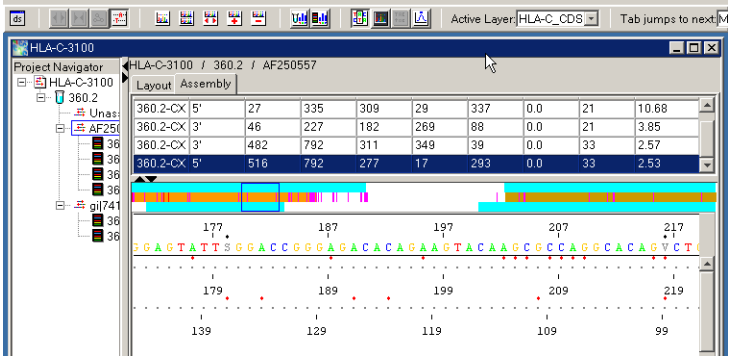
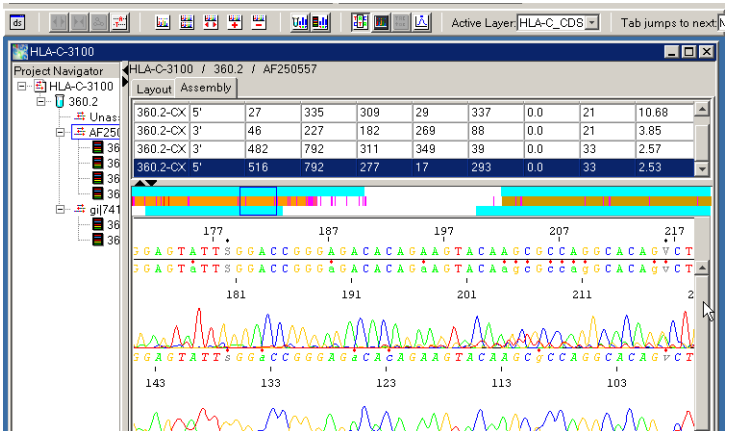
View	Procedure	Display
Sequence Assembly	Select the **Assembly** tab.	
Dot Assembly	1. Select the **Assembly** tab. 2. Click	
Electro-pherogram Assembly	1. Select the **Assembly** tab. 2. Select a sample in the sample table. 3. Click for multiple EPS or click for one EP.	

Table 7-2 Segment Views *(continued)*

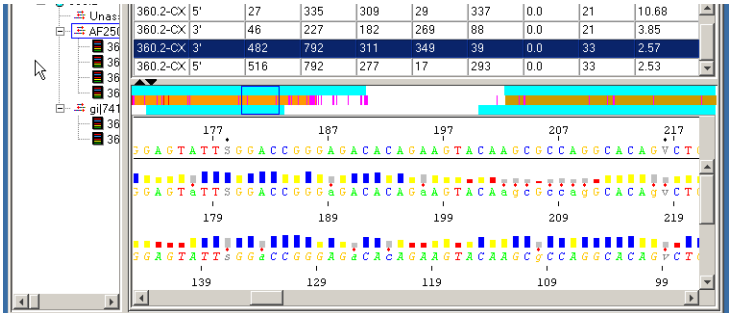
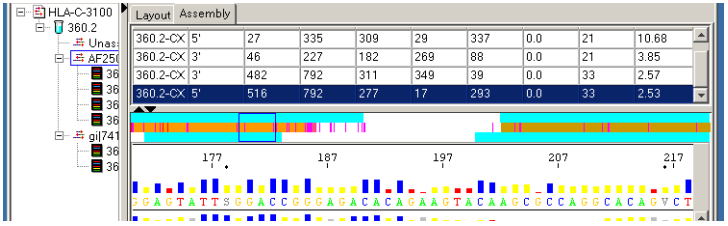
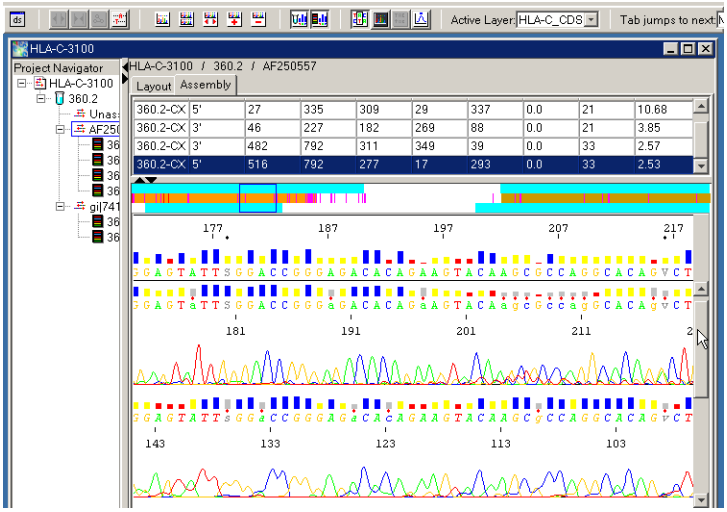
View	Procedure	Display
QV Assembly	1. Select the **Assembly** tab. 2. Click for sample QVs and/or click for consensus QVs.	 

Table 7-2 Segment Views *(continued)*

View	Procedure	Display
Assembly (The EP and QV views can be used with the Sequence or Dot view.)	1. Select the **Assembly** tab. 2. Click . 3. Click . 4. Click .	 **Note:** A black dot above a base in the consensus sequence indicates a discrepancy between the consensus and the reference sequences, and a red dot above a sample base indicates the base was edited by the consensus caller. Options: • Use (zoom tools) to zoom in/out the horizontal view. • Use (Aligned EP) to scale horizontal EP peaks to be of even width and aligned to base calls. • Use (View Column Selector) to turn on column selector, indicating position where base call is aligned to EP peak of original (not scaled) EP. • Use (Inverse Video) to switch between white and black background displays. • Use Find and Find Again to search for text.

Sample Views

The sample result includes all the data characteristics of a sample. Sample data characteristics are displayed in the following tabs:

Table 7-3 Sample View Tabs

Tab	Displayed Information
Annotation	Information about the data and its analysis.
Sequence	Sequence of the sample in NT codes. For readability, the display clusters the sequences into substrings of 10 characters each, separated by blanks.
Features	Calculated clear range and multiple base positions.
Electropherogram	Electropherogram and basecall data for the sample. The data excluded from the clear range is shown in gray.
Raw	Raw data collected by the genetic analyzer.

Displaying the Sample Views

To display sample views:

1. Open the project of interest.
2. In the navigation pane, open a specimen, then open a segment.
3. Select a sample, then select a tab (see Table 7-3).
4. Select a new tab to change the view.
5. To view a different sample, select a new sample from an open segment, then select a tab.

Examples of the five tab views are displayed in Table 7-4, “Sample Views,” on page 7-16.

Table 7-4 Sample Views

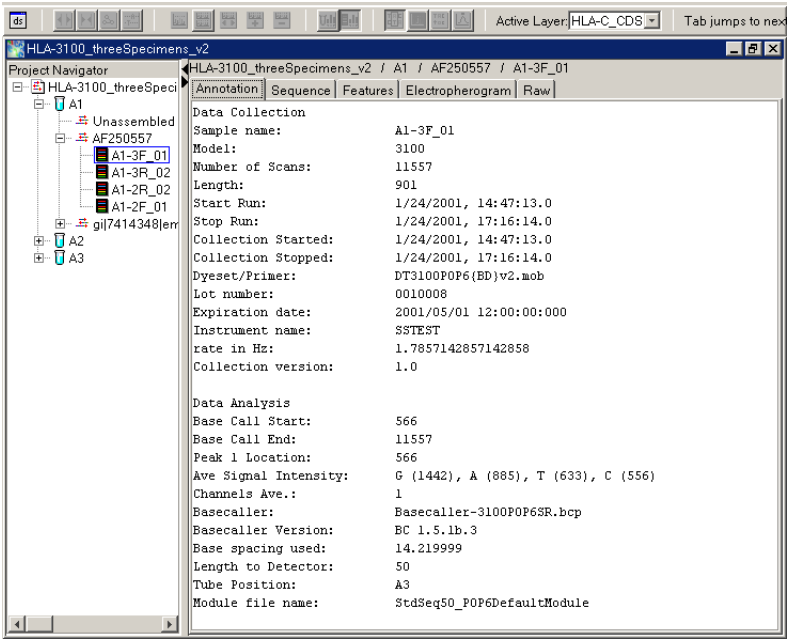
Tab	Display
Annotation	

Table 7-4 Sample Views *(continued)*

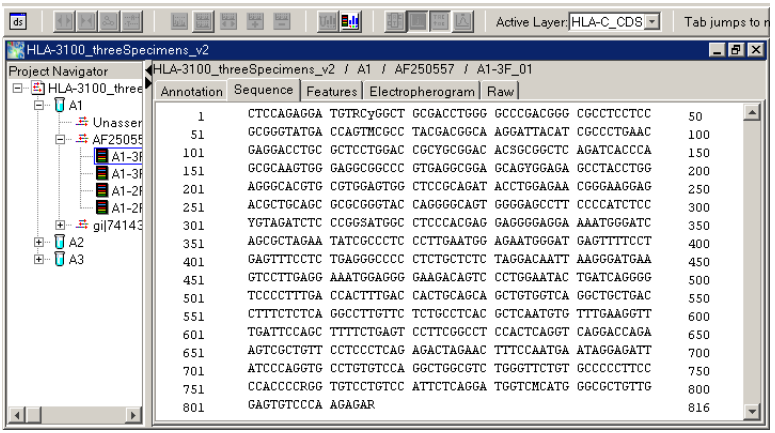
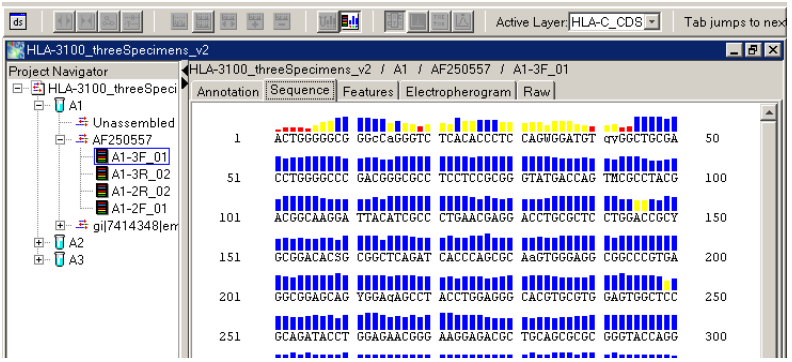
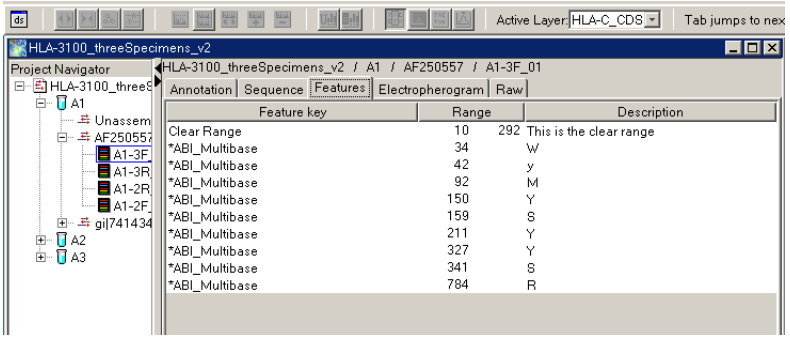
Tab	Display
Sequence	 Use (Show/Hide Sample QV) to display/hide sample QVs. 
Features	

Table 7-4 Sample Views *(continued)*

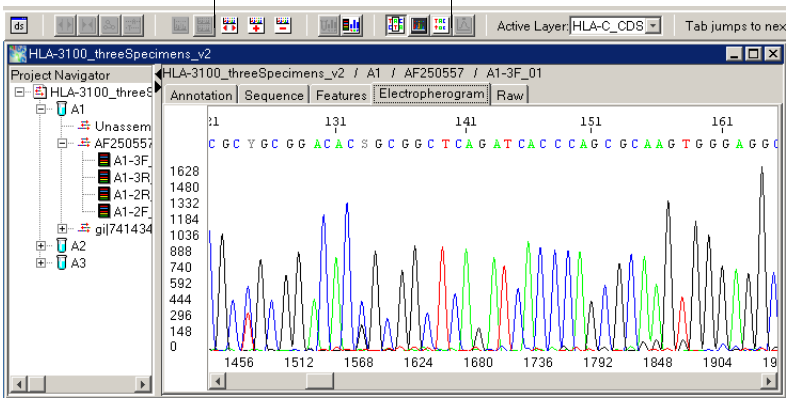
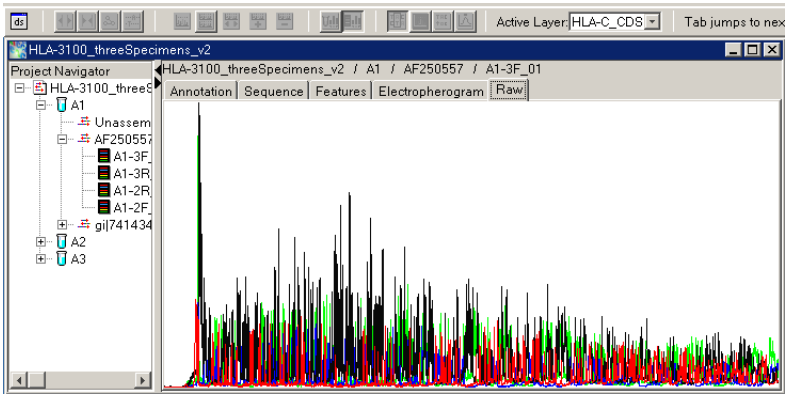
Tab	Display
Electropherogram	Active tools  Options: • Use (zoom tools) to zoom in/out the horizontal view. • Use (Show/Hide Sample QV) to display/hide sample QVs. • Use (Inverse Video) to switch between white and black background displays. • Use (View Original Sequence) to display/hide the original sequence called by the software. • Use Find and Find Again to search for text.

Table 7-4 Sample Views *(continued)*

Tab	Display
Raw	

Viewing Variants

Two methods to view variant data are presented here.

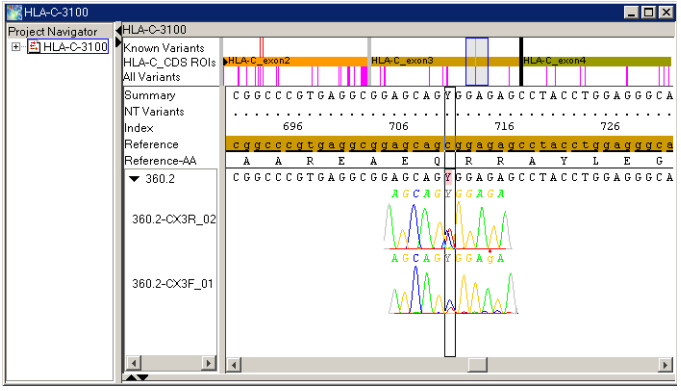
Note: To edit variant data, see “Editing Variants” on page 8-20.

To view variant data:

Method 1

1. Open the project of interest.
2. Click a consensus base.
3. View the electropherogram snippets by clicking the triangle that appears next to the specimen name.
4. In the Tab Jump to Next drop-down list, select **Multiple**, then select **Known Variant** and **Unknown Variant**.
5. Press **Tab** to move to the next variant or press **Shift-Tab** to move to the previous variant.

   Note: Pressing Ctrl+Z moves any electropherogram snippets of the selected variants to the middle of the view.



Method 2

1. Open the project of interest.
2. Select **Analysis > Report Manager** or click .
3. In the navigation pane, select the report you want to view.
4. Select **Window > Tile**.
5. Review the positions by selecting a base change in the Mutation table. This action brings the alignment view to the correct position in the alignment.

Saving Your Data

When you finish, select **File > Save Project** or click .

IMPORTANT! Any changes you make are saved and overwrite the existing project.

About the Reports

After the data is analyzed, you can view, export, and print reports. Reports can help you troubleshoot your results because reports contain hyperlinks to the primary sequence data. You can use reports with project results to evaluate your samples, modify the analysis settings, and edit the basecalling.

Types of Reports

Ten reports are generated with every project analysis. Each project has its own Report Manager window containing all of the following reports:

- Analysis QC
- Mutations
- AA Variants
- Specimen Statistics
- Sequence Confirmation
- Base Frequency
- Library Search
- RDG
- Audit Trail
- Genotyping

Note: Only one report can be viewed at a time.

Common to all reports is a Summary table that includes project information and the specimens in the report.

Exporting and Printing Reports

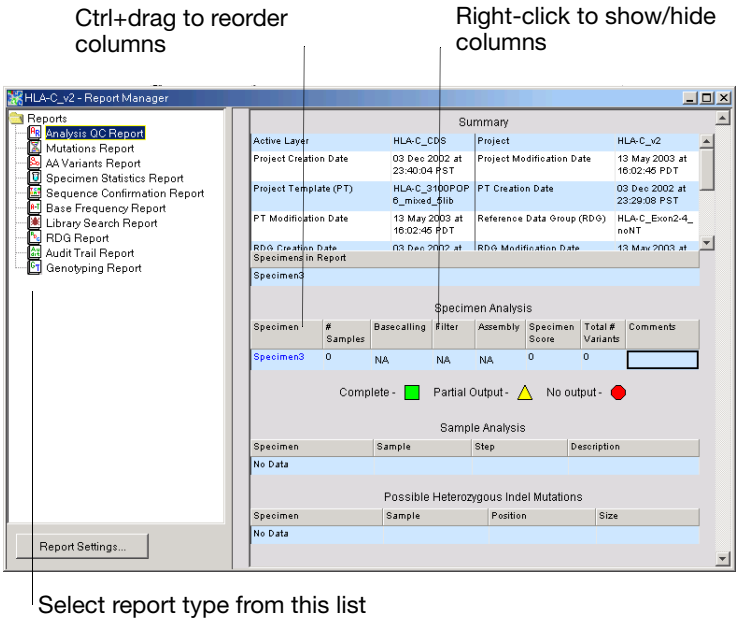
To export a report, see “Exporting Reports” on page 9-8, and to print a report, see “Printing Data and Reports” on page 9-11.

Viewing the Reports

The data in the reports is filtered, based on the view (project, specimen, segment, or sample) selected in the navigation pane of the project and the layer selected from the Active Layer drop-down menu.

To view a report:

1. Open the project of interest, then select the active layer.
2. In the navigation pane, select the project, specimen, segment, or sample view.
3. Select **Tools** > **Report Manager** or click .



4. In the navigation pane, select the report you want to view.
5. To view other reports, select a different report in the navigation pane.

   Note: Only one report can be viewed at a time.

6. To update the reports with additional data:
   a. In the project, select additional or different specimen/samples.
   b. Click to update the data in the open report.

Viewing the Reports and Project Results

When you view reports, you can tile the report with the Project window so that you can easily view the data when you use a hyperlink.

To view the project results and reports together:

1. Open the project of interest.
2. Select a layer in the Active Layer drop-down list.
3. Select **Analysis** > **Report Manager** or click .
4. In the navigation pane, select the report you want to view.
5. Select **Window** > **Tile**.
6. Click a hyperlink (blue text) in the report, then view the data in the Project view.

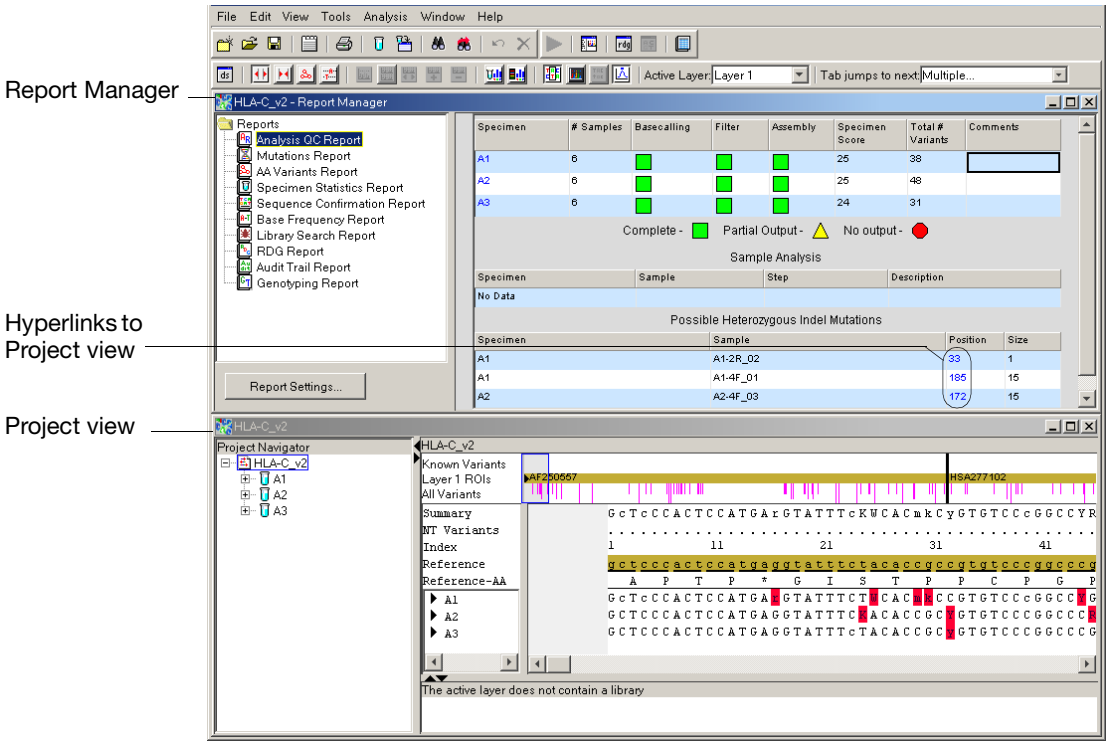


Figure 7-2 Viewing Results and Reports Together

Analysis QC Report

Applied Biosystems recommends that you click before viewing the Analysis QC report.

This report contains four separate tables. All blue text is hyperlinked to the project navigator.

Table 7-5 Parts of the Analysis QC Report

Table	Description
Summary	Displays project information and the specimens in the report.
Specimen Analysis	Displays specimen analysis results, specimen score (average consensus QV) and total number of variants.
Sample Analysis	Displays sample analysis errors and details.
Possible Heterozygous Indel Mutations	Displays possible mutations and their positions and size for each specimen.

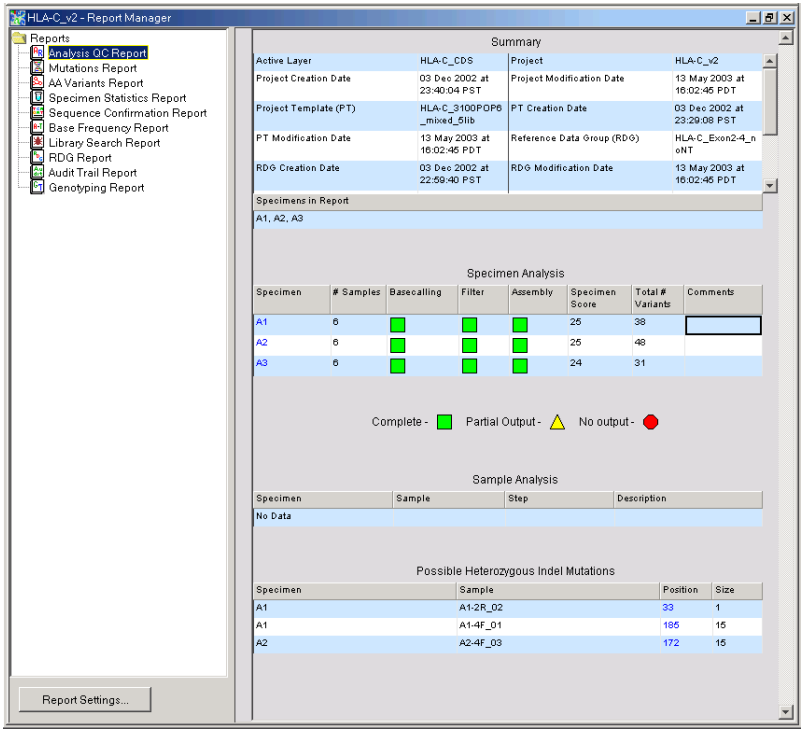


Figure 7-3 Analysis QC Report

Mutations Report

The report contains two separate tables. All blue text is hyperlinked to the project navigator.

Table 7-6 Parts of the Mutation Report

Table	Description
Summary	Displays project information and the specimens in the report.
Mutation	Displays the bases changed, ROI, position, length, type, QV, and effect information for each mutation detected in a specimen.

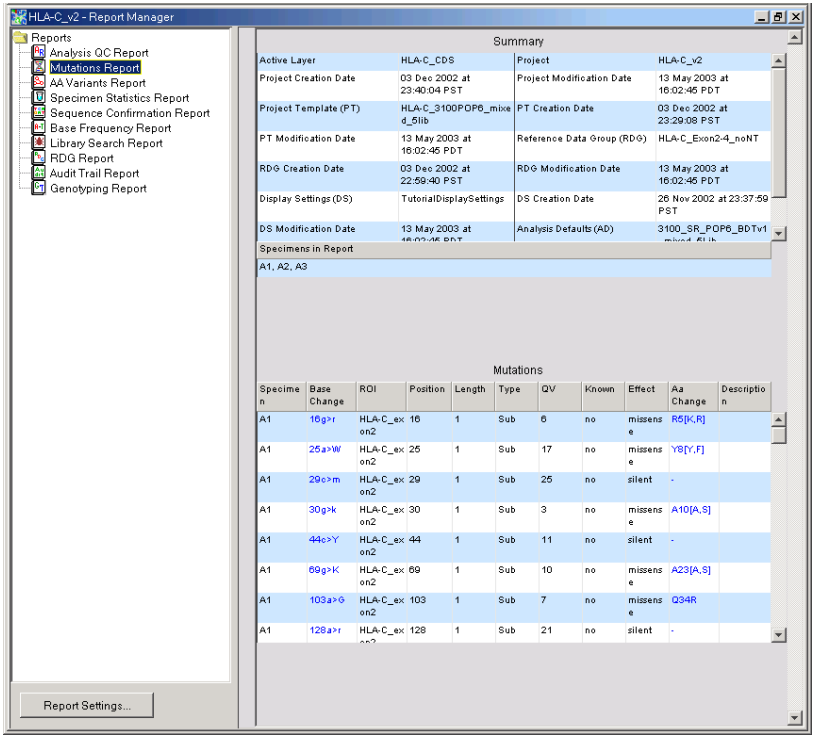


Figure 7-4 Mutations Report

The Mutations report includes a column that provides a predicted "effect" for each nucleotide variant. Table 7-7 describes the possible values in the Effect column.

In the Mutation report, clicking an NT variant links to the corresponding base in the Project View. The corresponding AA change in the Mutation Report links to the AA Variant in the AA Variants Report. This AA variant is, in turn, linked to the AA in the Project AA view.

Table 7-7 Predicted Effects of Nucleotide Variants

Effect	Description
Missense	The substitution variant codes for an amino acid substitution.
Nonsense	The substitution variant codes for a terminator codon. (In a mixed codon, if any codon is a terminator codon "nonsense" will be displayed).
Silent	The substitution variant is in a coding region but does not code for an amino acid change.
Frameshift insertion	The insertion variant is in a coding region and codes for a frameshift in translation (the size of the insertion is not a multiple of three).
Frameshift deletion	The deletion variant is in a coding region and codes for a frameshift in translation (the size of the deletion is not a multiple of three).
In-frame insertion	The insertion variant is in a coding region and does not code for a frameshift in translation (the size of the insertion is a multiple of three).
In-frame deletion	The deletion variant is in a coding region and does not code for a frameshift in translation (the size of the deletion is a multiple of three).
Non-coding	The variant is not in a coding region.
Partial codon	The variant is in a coding region, but occurs at the beginning or end of the sequence, where you do not know the full three-base codon sequence.

Table 7-7 Predicted Effects of Nucleotide Variants

No information	The variant is a result of the consensus sequence not completely covering the reference sequence. These are not real variants, so you cannot predict a real effect.
Heterozygous deletion/insertion	Specimen-level heterozygous indel mutation (HIM) identification

AA Variants Report

The report contains two separate tables. All blue text is hyperlinked to the project navigator.

Table 7-8 Parts of the AA Variant Report

Table	Description
Summary	Displays project information and the specimens in the report.
AA Variant	Displays the AA changed, position, length, type, and description for each variant detected in a specimen.

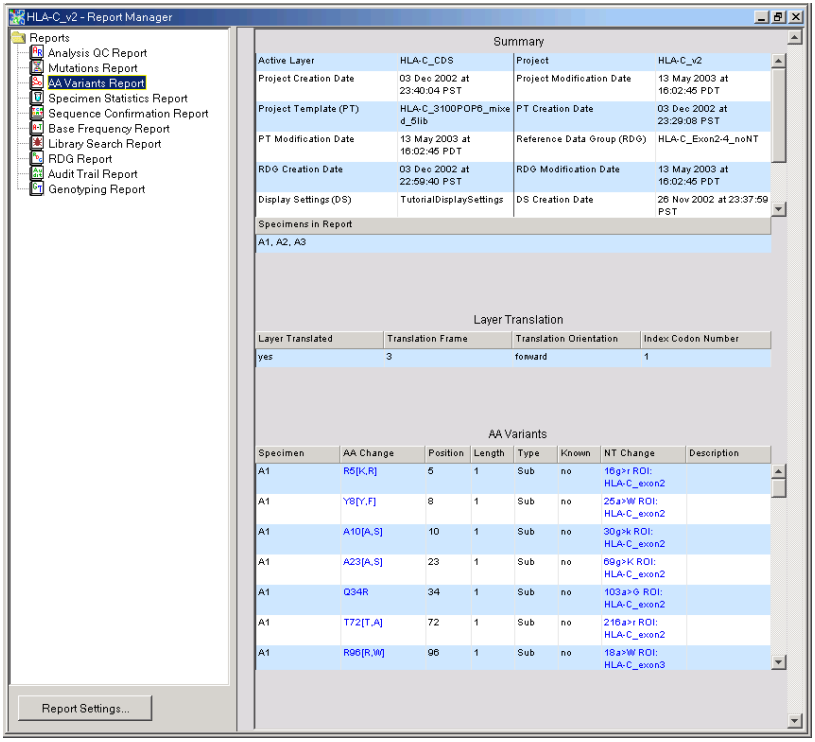


Figure 7-5 AA Variant Report

Specimen Statistics

The report contains three separate tables.

Table 7-9 Parts of the Specimen Statistics Report

Table	Description
Summary	Displays project information and the specimens in the report.
Specimen Statistics	Displays the bases changed, ROI, position, length, type, QV, and effect information for each mutation detected for each specimen.
Sample Results	Displays the specimen, segment, assembly status, calculated clear range, sample score (average QV) and% mixed bases for each sample.

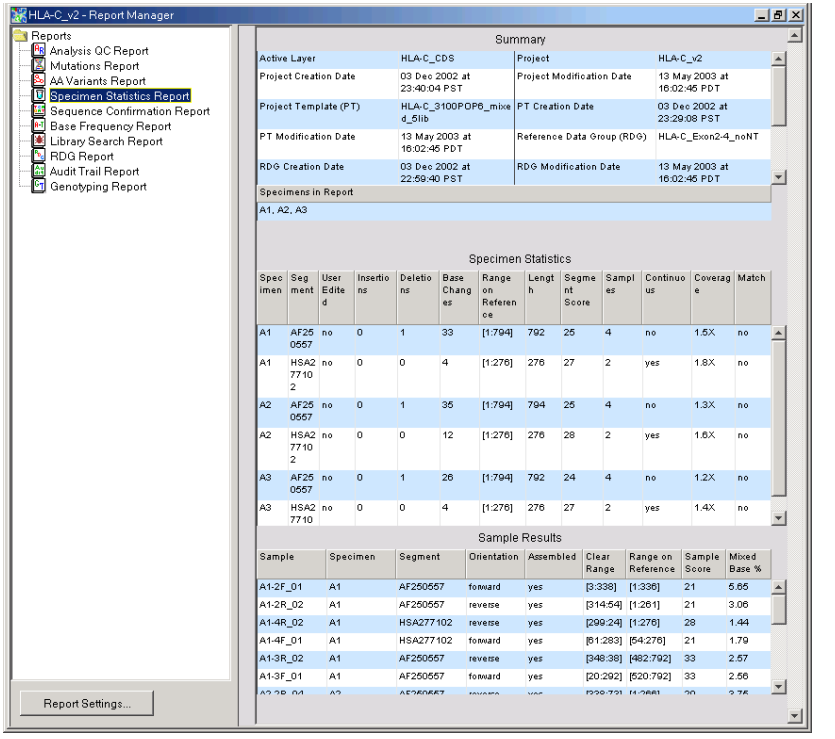


Figure 7-6 Specimen Statistics Report

Sequence Confirmation Report

The report contains two separate tables.

Table 7-10 Parts of the Sequence Confirmation Report

Table	Description
Summary	Displays project information and the specimens in the report.
Sequence Confirmation	Displays the match, the number of insertions, deletions and bases, and the amount of coverage and whether it is continuous for each specimen.

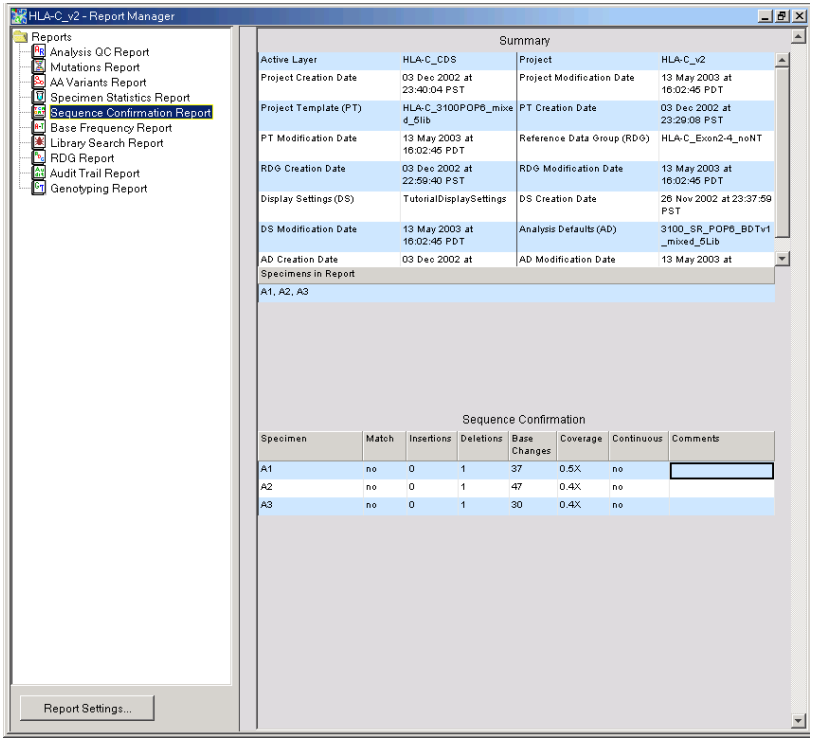


Figure 7-7 Sequence Confirmation Report

Base Frequency Report

The report contains two separate tables.

Table 7-11 Parts of the Base Frequency Report

Table	Description
Summary	Displays project information and the specimens in the report.
Base Frequency	Displays the reference, ROI, and the% of each base and space for each variant position.

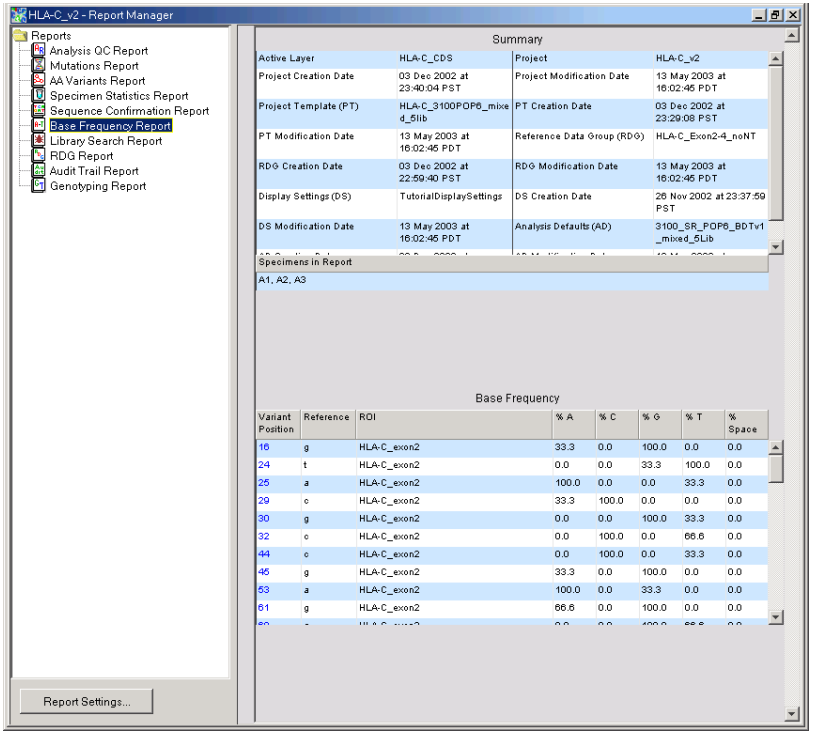


Figure 7-8 Base Frequency Report

Library Search Report

The report contains five separate tables (see Figure 7-9 on page 7-36).

Table 7-12 Parts of the Library Search Report

Table	Description
Summary	Displays project information and the specimens in the report.
Library	Displays the library and information used in the search.
Specimen Results	Displays the match status, crucial position and constant position errors for each specimen.
Hit List	Displays the library matches found, their scores (closest match), and mismatch information for each specimen.
Constant Positions Errors	Displays the position, specimen base and library base for each specimen and ROI. The values in the Position column are hyperlinked to the project navigator.

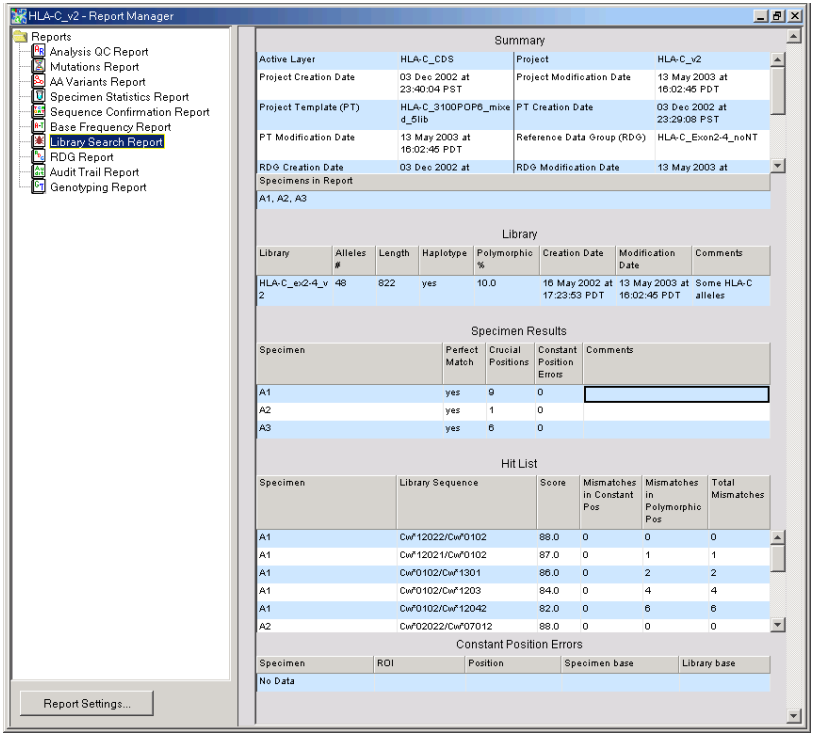


Figure 7-9 Library Search Report

RDG Report

The report contains three separate tables.

Table 7-13 Parts of the RDG Report

Table	Description
Summary	Displays project information and the specimens in the report.
Layers	Displays a summary of the information for each layer in the project as defined in the RDG.
ROIs	Displays a summary of the information for each ROI as defined in the RDG.

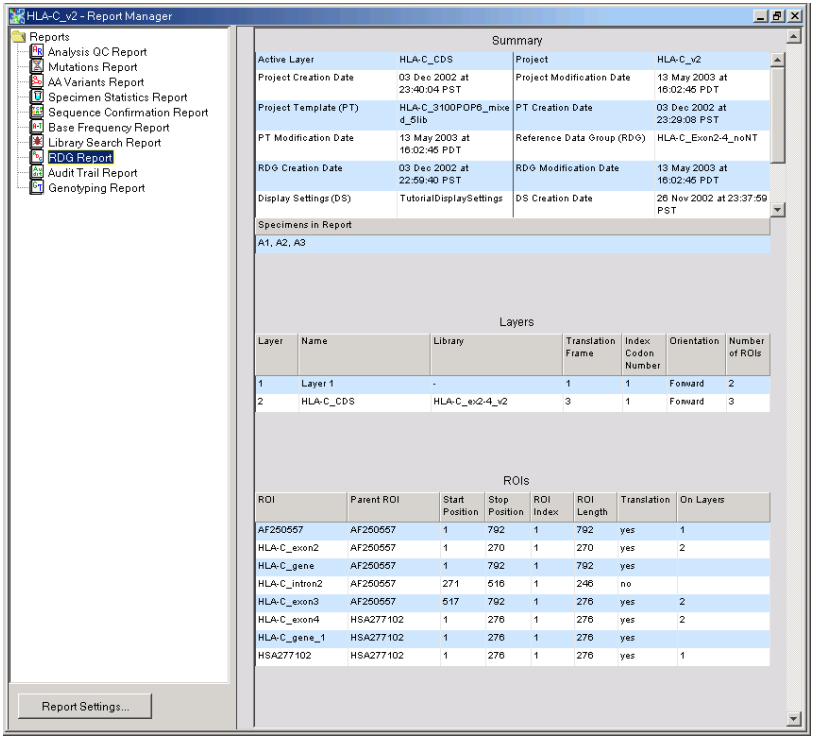

HLA-C_v2 - Report Manager
Reports
Analysis QC Report
Mutations Report
AA Variants Report
Specimen Statistics Report
Sequence Confirmation Report
Base Frequency Report
Library Search Report
RDG Report
Audit Trail Report
Genotyping Report
Report Settings...
Summary
Active Layer HLA-C_CDS
Project HLA-C_v2
Project Creation Date 03 Dec 2002 at 23:40:04 PST
Project Modification Date 13 May 2003 at 16:02:46 PDT
Project Template (PT) HLA-C_3100POP6_mixed_5lib
PT Creation Date 03 Dec 2002 at 23:29:08 PST
PT Modification Date 13 May 2003 at 16:02:46 PDT
Reference Data Group (RDG) HLA-C_Exon2-4_noNT
RDG Creation Date 03 Dec 2002 at 22:59:40 PST
RDG Modification Date 13 May 2003 at 16:02:46 PDT
Display Settings (DS) TutorialDisplaySettings
DS Creation Date 26 Nov 2002 at 23:37:59 PST
Specimens in Report
A1, A2, A3
Layers
Layer Name Library Translation Frame Index Codon Number Orientation Number of ROIs
1 Layer 1 - 1 1 Forward 2
2 HLA-C_CDS HLA-C_ex2-4_v2 3 1 Forward 3
ROIs
ROI Parent ROI Start Position Stop Position ROI Index ROI Length Translation On Layers
AF250557 AF250557 1 792 1 792 yes 1
HLA-C_exon2 AF250557 1 270 1 270 yes 2
HLA-C_gene AF250557 1 792 1 792 yes
HLA-C_intron2 AF250557 271 516 1 246 no
HLA-C_exon3 AF250557 517 792 1 276 yes 2
HLA-C_exon4 HSA277102 1 276 1 276 yes 2
HLA-C_gene_1 HSA277102 1 276 1 276 yes
HSA277102 HSA277102 1 276 1 276 yes 1

Figure 7-10 RDG Report

Audit Trail Report

The report contains two separate tables.

Table 7-14 Parts of the Audit Trail Report

Table	Description
Summary	Displays project information and the specimens and samples in the report.
Audit Trail	Displays a record of the edits and changes made to data in a project, if the audit trail feature is on.

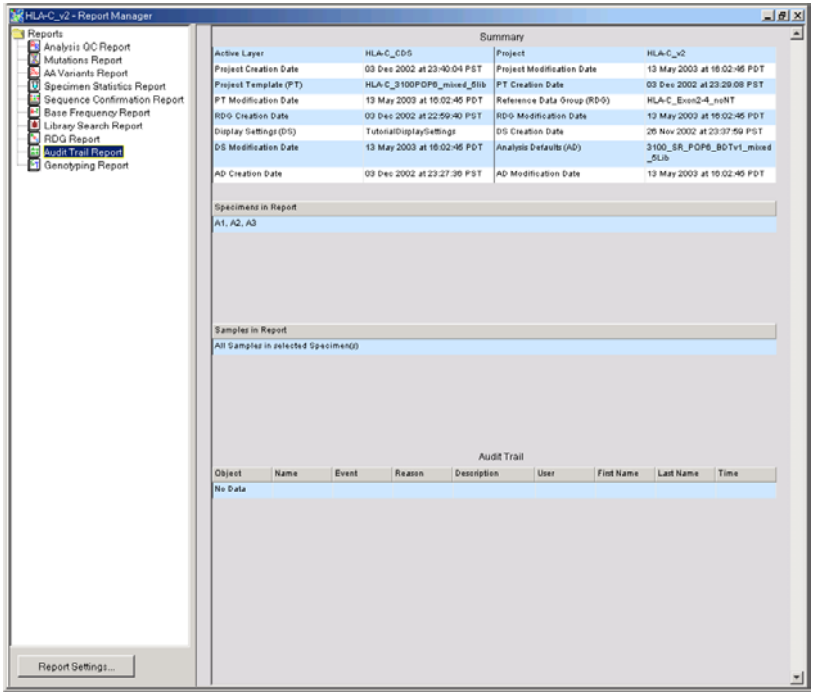


Figure 7-11 Audit Trail Report

Genotyping Report

The Genotyping report contains five separate tables.

Table 7-15 Parts of the Genotyping Report

Table	Description
Summary	Displays project information and the specimens and samples in the report
Specimens_in_ Report	Lists the specimens included in the report
Gene Summary	These fields are not used at this time
Transcript Table	Displays the layout of the intended amplicons. Lists all of the transcripts with their accompanying amplicons in the RDG. • Layer: Name of the transcript being reported on • Amplicon_ID: Identity of an amplicon on this transcript (there are many rows of amplicons for each transcript) • Assay_Target_Start: Start of the assay region of the amplicon in reference segment coordinates • Assay_Target_ Length: Designed length of the assay region • 5'_Regulatory_Length: Amount of amplicon sequence covering 5' regulatory region • Exon_Length: Amount of amplicon sequence covering exon region in this transcript • Intron_Length: Amount of amplicon sequence covering intron region in this transcript

Table 7-15 Parts of the Genotyping Report

Table	Description
Resequencing Coverage	Displays specimen coverage statistics for the active layer. • Specimen: Name of specimen being reported on • FGR Coverage: Full gene region coverage. Percent of specimen consensus sequence covering the entire targeted region • FGR Avg QV: Average quality value of the consensus sequence across the target region • FGR Fwd Coverage: Percent of target region coverage by specimen consensus sequence generated from forward reads • FGR Rev Coverage: Percent of target region coverage by specimen consensus sequence generated from reverse reads • 5' Reg Coverage: 5' regulatory region coverage. Percent of specimen consensus sequence covering the 5' regulatory region • 5' Reg Coverage Avg QV: Average quality value of the consensus sequence at the 5' region • 5' Reg Fwd Coverage: Percent of target region coverage by specimen consensus sequence generated at the 5' region for forward reads • 5' Reg Rev Coverage: Percent of target region coverage by specimen consensus sequence generated at the 5' region for reverse reads • Exon Coverage: Exon coverage for a specific gene • Exon Avg QV: Average quality value of the consensus sequence at the exon region • Exon Fwd Coverage: Quality value of the consensus sequence at the exon region for forward reads • Exon Rev Coverage: Quality value of the consensus sequence at the exon region for reverse reads Intron Coverage: Intron coverage for a specific gene • Intron Avg QV: Average quality value of the consensus sequence at the intron region • Intron Fwd Coverage: Quality value of the consensus sequence at the intron region for forward reads • Intron Rev Coverage: Quality value of the consensus sequence at the intron region for reverse reads
Genotyping Table	Displays genotypes of all specimens at selected positions in the reference.

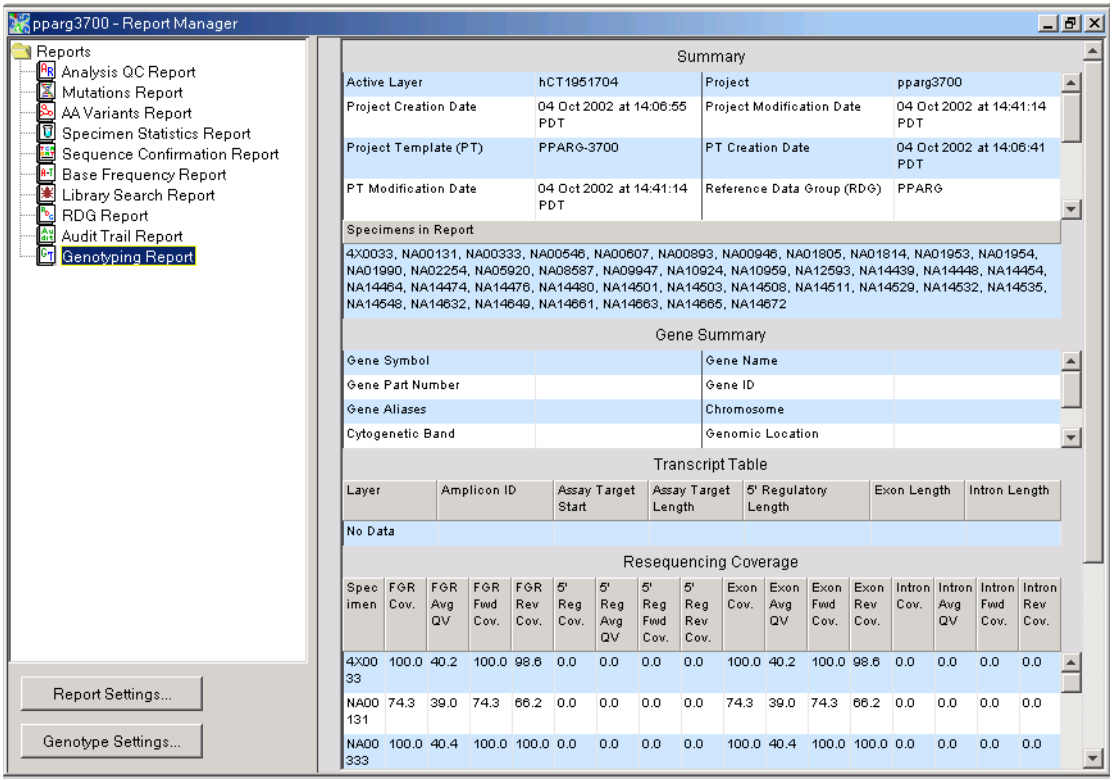


Figure 7-12 Genotyping Report

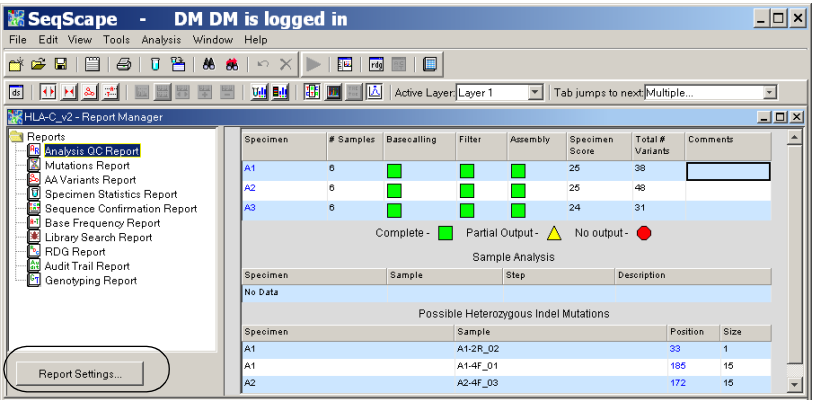
Customizing the Reports

Customizing Text Settings

You can customize the text and select the amino acid variant.

To customize the report settings:

1. Click the **Report Settings** button on the Report Manager window.



The Report Settings dialog box opens displaying the Report Display Settings tab.

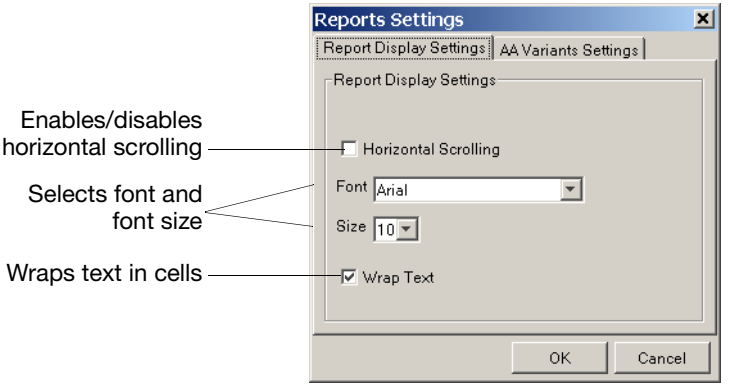


Figure 7-13 Report Display Settings tab of the Report Settings dialog box

2. Select or deselect **Horizontal Scrolling** check box. The default is off.
3. Select a font type and font size in the appropriate drop-down lists. The default is font is Arial size 10.
4. Select or deselect **Wrap Text**. Examples of wrapped and unwrapped text are shown in Figure 7-14.

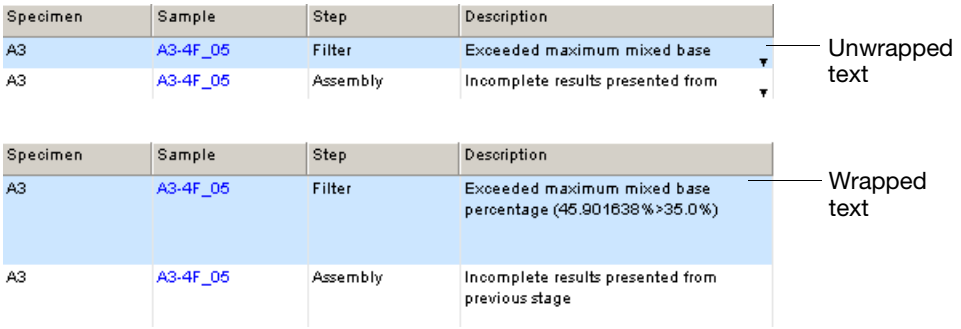
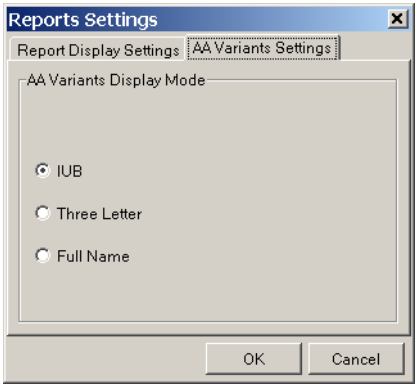


Figure 7-14 Examples of unwrapped and wrapped text

To set the display of AA variants:

1. Click the **AA Variants Settings** tab.

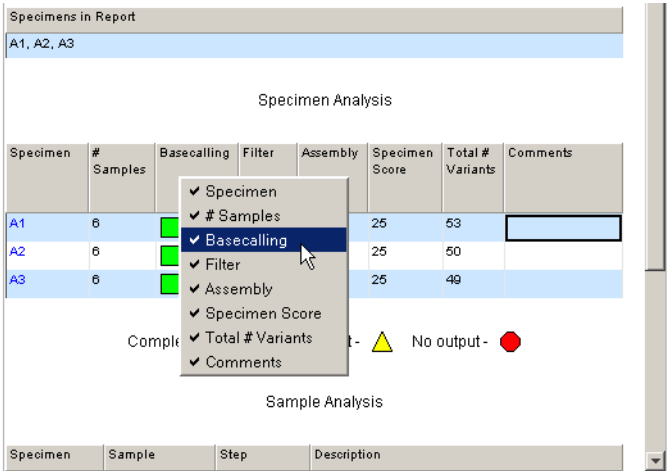


2. Select display mode then click **OK**.

Customizing the Data View

To customize the information displayed in the report:

1. Right-click any column heading of a table. A list of the column headings in the table is displayed.



2. To hide a column, deselect the column heading.
3. Repeat steps 1 and 2 to deselect additional headings.
4. To redisplay a column, right-click any column heading, then select the column heading.
5. To sort the data A to Z or Z to A in a Sample Details or Errors table column, double-click the column heading. Double-click again to sort in the opposite direction.
6. To customize the table header and footer information, see “Customizing Header and Footer Display” on page 9-11.

Reanalyzing and Editing Data 8

This chapter contains:

Workflow for This Chapter 8-2
About Analysis Parameters 8-3
Changing the Analysis Parameters in the Sample Manager 8-6
Changing the Analysis Parameters in an Analysis Protocol 8-7
Editing the Data. 8-13
Editing a Sample or a Consensus Sequence 8-14
Adjusting the Clear Range 8-16
Editing Variants 8-20
Adding a Genotype 8-22

Workflow for This Chapter

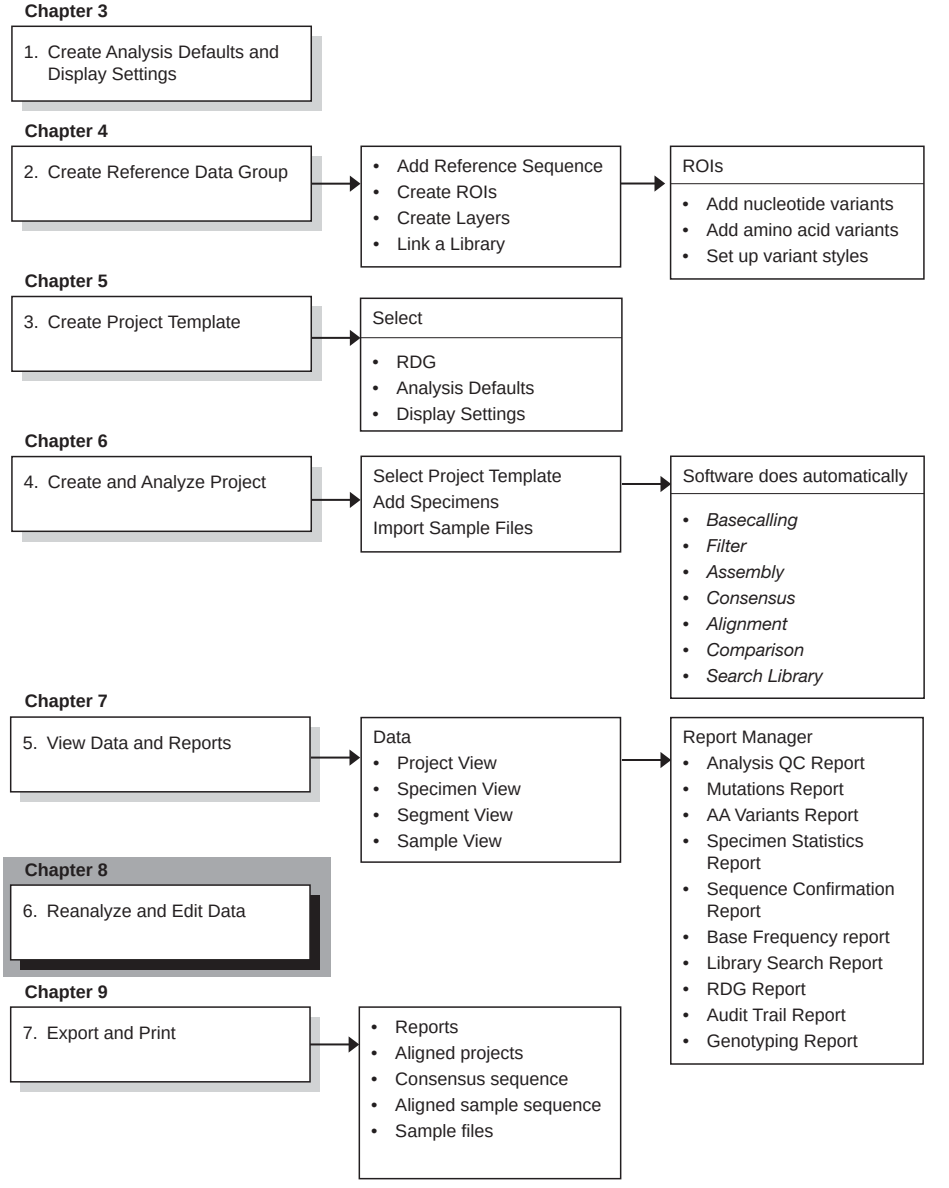


Figure 8-1 Reanalyzing and Editing Data Step

About Analysis Parameters

Introduction

The analysis parameters (basecaller and DyeSet/Primer file) associated with every sample file are used when the sample files are analyzed.

Sometimes poor project results can be corrected or improved by changing certain analysis settings and applying the new settings to the affected samples.

Common examples of errors that affect basecalling are:

- Incorrect stop point selected
- Bad base spacing
- Poor quality data
- Incorrect basecaller and/or dyeset/primer used for basecalling
- Wrong peak 1 location and start point calculated by the software

Note: Refer to the *ABI PRISM® DNA Sequencing Analysis Software User Guide* for instructions to define a new peak 1 and the start and stop point locations.

Viewing Analysis Parameters in the Sample Manager

You can use the Sample Manager to display sample files and their current analysis information including the basecaller and DyeSet/Primer files (see Figure 8-2 on page 8-4). The analysis parameters can be modified and applied to samples. You can apply these changes to one sample, some samples, or all samples in the Sample Manager.

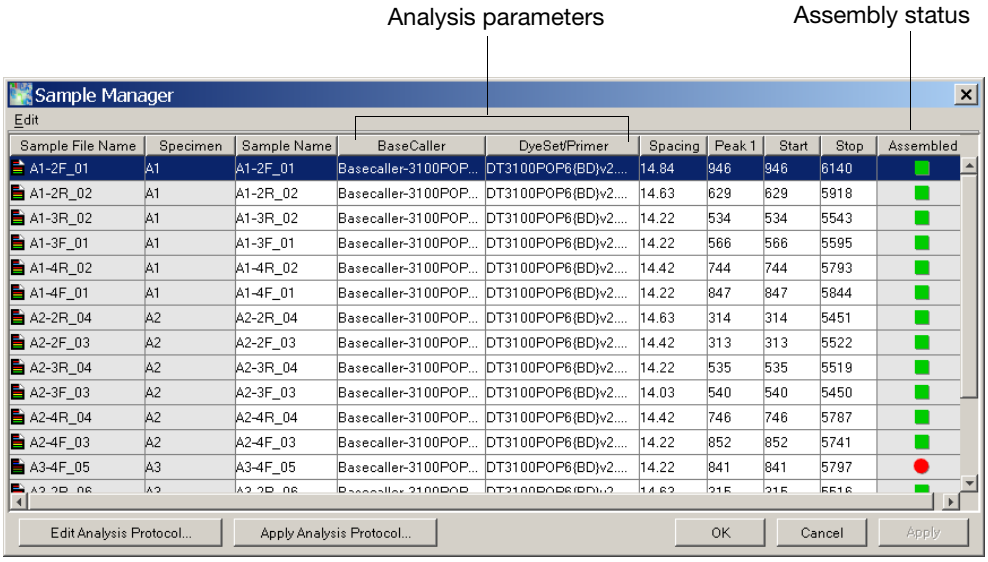

Analysis parameters
Assembly status
Sample Manager
Edit
Sample File Name
Specimen
Sample Name
BaseCaller
DyeSet/Primer
Spacing
Peak 1
Start
Stop
Assembled
A1-2F_01 A1 A1-2F_01 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.84 946 946 6140
A1-2R_02 A1 A1-2R_02 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.63 629 629 5918
A1-3R_02 A1 A1-3R_02 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.22 534 534 5543
A1-3F_01 A1 A1-3F_01 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.22 566 566 5595
A1-4R_02 A1 A1-4R_02 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.42 744 744 5793
A1-4F_01 A1 A1-4F_01 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.22 847 847 5844
A2-2R_04 A2 A2-2R_04 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.63 314 314 5451
A2-2F_03 A2 A2-2F_03 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.42 313 313 5522
A2-3R_04 A2 A2-3R_04 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.22 535 535 5519
A2-3F_03 A2 A2-3F_03 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.03 540 540 5450
A2-4R_04 A2 A2-4R_04 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.42 746 746 5787
A2-4F_03 A2 A2-4F_03 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.22 852 852 5741
A3-4F_05 A3 A3-4F_05 Basecaller-3100POP... DT3100POP6{BD}v2.... 14.22 841 841 5797
Edit Analysis Protocol...
Apply Analysis Protocol...
OK
Cancel
Apply

Figure 8-2 Sample Files in the Sample Manager

The Sample Manager window displays the following information:

Table 8-1 Information in the Sample Manager

Column Heading	Description
Sample File Name	Information from the plate record and project. It cannot be changed in the Sample Manager.
Specimen	Information from the plate record and project. It cannot be changed in the Sample Manager.
Sample Name	Name of the sample, taken from the plate record. It can be changed.
Basecaller	Algorithm used to call the bases. It can be changed.

Table 8-1 Information in the Sample Manager *(continued)*

Column Heading	Description
DyeSet/Primer	A DyeSet/Primer file corrects for mobility shifts and color-code changes, depending on which chemistry was used. DyeSet/Primer files are sometimes known as mobility or .mob files. All mobility files have the extension .mob. It can be changed.
Spacing	The number of scan points from the crest of one peak to the crest of the next peak. During basecalling, a spacing calibration curve is applied to the data to determine a base spacing value.
Peak 1	The first data point that is from the sample, not including primer peaks in dye primer chemistries. It is the reference point for the spacing and mobility corrections performed by the basecalling software.
Start	The raw data point where the basecalling starts in the sample file. The Start Point is normally the same as the beginning of the first base peak.
Stop	Specifies the last raw data point to be included in the basecalling. If the default Stop Point is used, this endpoint is the last data point in the file.
Assembled	Displays the assembly status of the sample. A green box means assembled and a red circle means not assembled.

Changing the Analysis Parameters in the Sample Manager

Adding Samples to the Sample Manager

To add samples to the Sample Manager:

1. Open the project of interest.
2. Select a layer in the Active Layer drop-down list.
3. In the navigation pane:

To add all …	Select the …
Samples in a project	Project icon
Samples in a specimen	Specimen
Selected samples in a segment	Segment
Selected sample(s)	Sample(s)*

* Use the Shift key to select contiguous samples, or use the Ctrl key to select noncontiguous samples.

4. Select **Analysis** > **Sample Manager**.

   The selected files are displayed in the Sample Manager.

Changing Basecaller and DyeSet/Primer Files

Note: Use the basecaller and DyeSet/Primer tables in Appendix B to select the correct combination of files.

To change the basecaller and/or DyeSet/Primer file:

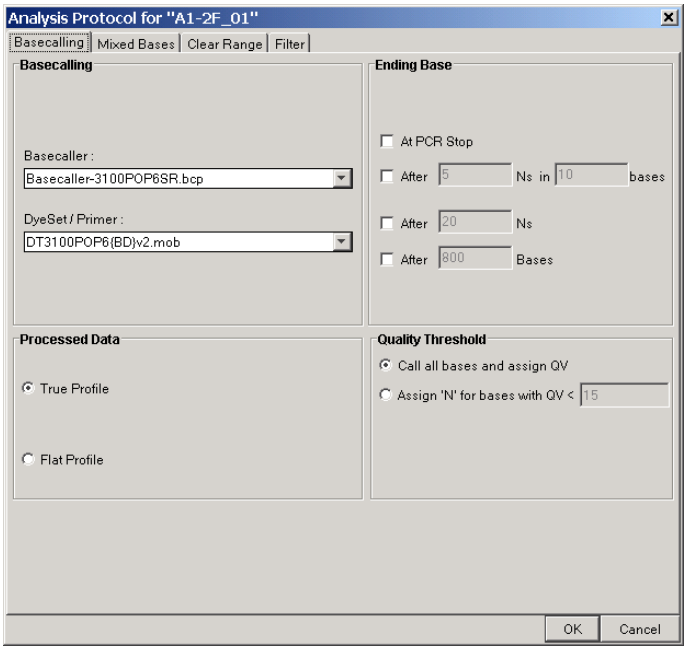
1. In the Sample Manager, select the sample you want to change.
2. In the Basecaller drop-down list, select a new basecaller.
3. In the DyeSet/Primer drop-down list, select a new DyeSet/Primer file.
4. To change multiple samples, use the Fill Down function.
5. Click **Apply**.
6. Click **OK**.
7. Click .

Changing the Analysis Parameters in an Analysis Protocol

Editing an Analysis Protocol

To edit an analysis protocol:

1. Add the samples to the Sample Manager (see “Adding Samples to the Sample Manager” on page 8-6).
2. In the Sample Manager window, click **Edit Analysis Protocol**.



3. In the **Basecalling** tab:
   a. In the **Basecalling** group, select the correct Basecaller and DyeSet/Primer files in the drop-down lists.
   b. In the **Editing Base** group, select one or more stop points for data analysis (optional).

c. In the **Processed Data** group, select either:

True Profile – Scales the processed traces uniformly, so that the average height of the peaks in the region of strongest signal is about equal to a fixed value (*e.g.*, 1000). The profile of the processed traces will be very similar to that of the raw traces.

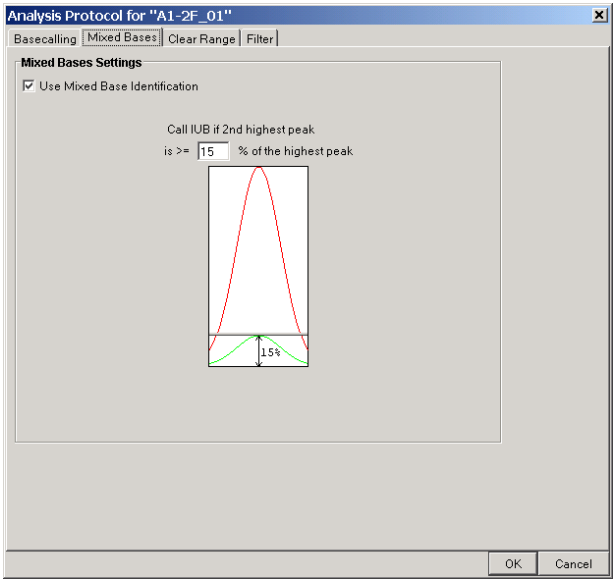
Flat Profile – Scales the processed traces semi-locally so that the average height of peaks in any region is about equal to a fixed value (*e.g.*, 1000). The profile of the processed traces will be flat on an intermediate scale (> about 40 bases).

d. In the **Quality Threshold** group, select either:

Call all bases and assign QV – KB Basecaller calls all bases and assigns a QV.

Assign 'N' for bases with QV < x – KB Basecaller calls Ns for all the bases that have QVs less than the value you enter.

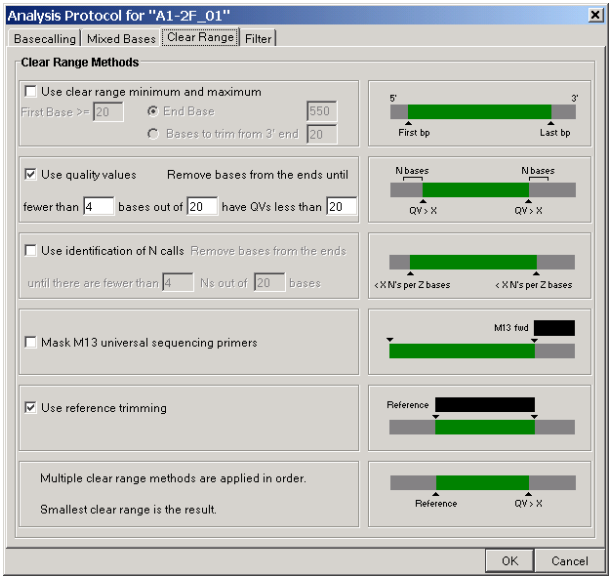
4. Select the **Mixed Bases** tab.



   a. If desired, select **Use Mixed Base Identification**.

   b. Use the default detection level of 25% or change the it by entering a new value or by dragging the % line up or down.

5. Select the **ClearRange** tab.

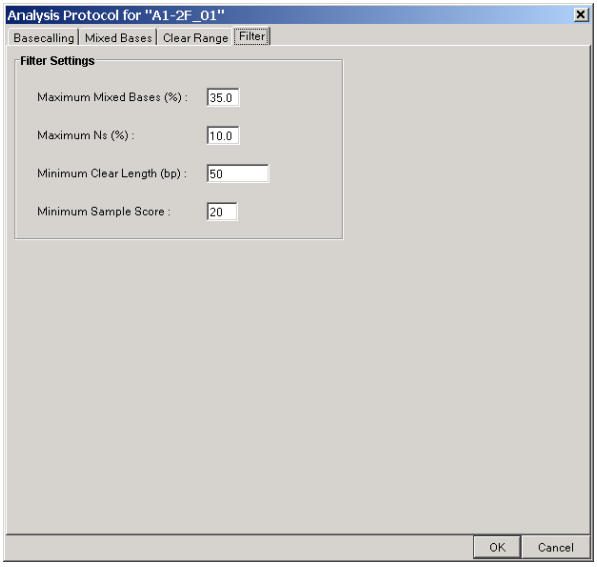


a. If desired, select one or more stop points for data analysis.

b. If desired, select to mask M13 primers from the clear range.

c. Select **Use reference trimming**.

6. Select the **Filter** tab.



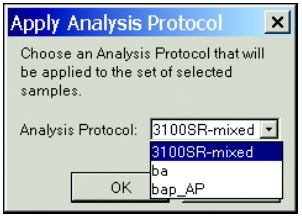
   Specify the values for the filter Settings.

7. Click **OK** to save the protocol and close the Analysis Protocol dialog box.

Applying the Analysis Protocol

To apply an analysis protocol:

1. Select the samples in the Sample Manager to apply the new settings to.
2. Click **Apply Analysis Protocol**.



3. Select a protocol from the Analysis Protocol drop-down list.
4. Click **OK**.

   The spacing, peak 1, and start and stop points change to zero.
5. Click **Apply**.

   The Assembled indicator changes from green (assembled) to red (unassembled), and the Analysis button becomes active.
6. Click .

Editing the Data

About Sequence Editing

To edit a sequence, you can:

- Adjust the clear range
- Add, delete, or change a base in a sample
- Add or delete a space in a sample
- Add, delete, or change a base in a specimen consensus
- Add or delete a space in a specimen consensus
- Add or delete a space in a reference

You can edit sequences within a project. The change is immediately reflected in the consensus sequence. You can also edit the consensus sequence. In this case, all the samples change to reflect the consensus edits. You can edit consensus sequences when viewing the data in the Specimen view or in the Project view.

Note: An edited base change or insertion appears in lowercase to distinguish it from an unedited base. This applies to both user edits and consensus-caller edits. See “Editing Bases with Quality Values” on page 10-10 for more information on editing bases with QVs.

When to Edit the Data

After analysis is complete, and you generate the analysis reports, depending on the results in the reports, you may want to:

- Adjust the clear range for a sample (see “Adjusting the Clear Range” on page 8-16)
- Edit a base or space in a sample or specimen (see “Editing a Sample or a Consensus Sequence” on page 8-14)

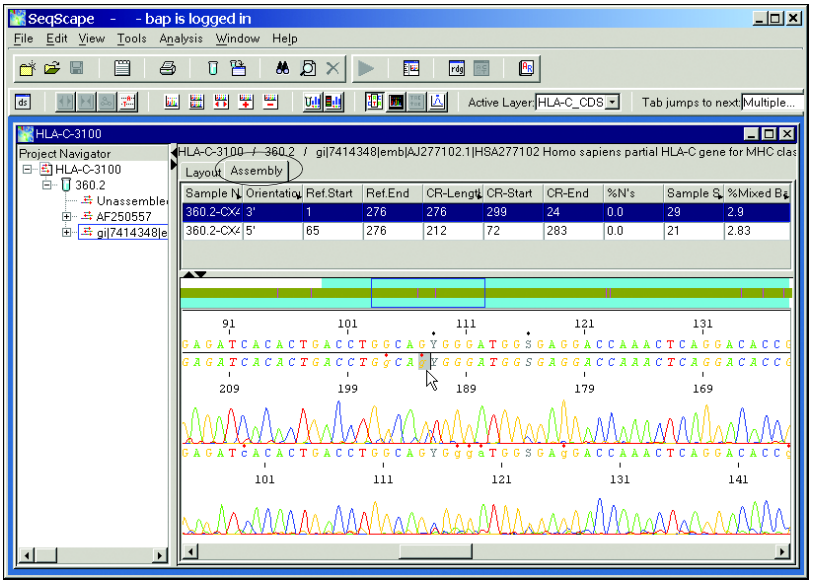
Editing a Sample or a Consensus Sequence

Editing a Consensus Sequence in the Segment View

To insert, delete, or change a base in the Specimen view:

1. Select the Segment icon in the navigation pane in the Project window.

   The Specimen view opens in the project document window.


SeqScape - bap is logged in
File Edit View Tools Analysis Window Help
Active Layer HLA-C_CDS
Tab jumps to next Multiple...
HLA-C-3100
Project Navigator
HLA-C-3100
360.2
Unassembled
AF250557
gi|7414348|e
HLA-C-3100 / 360.2 / gi|7414348|emb|AJ277102.1|HSA277102 Homo sapiens partial HLA-C gene for MHC clas
Layout Assembly
Sample N Orientatio Ref.Start Ref.End CR-Length CR-Start CR-End %N's Sample S %Mixed Ba
360.2-CX4 3' 1 276 276 299 24 0.0 29 2.9
360.2-CX4 5' 65 276 212 72 283 0.0 21 2.83

2. Select the **Assembly** tab, then select a layer in the Active Layer drop-down list.
3. To change or delete a consensus base, click the base you want to edit, then delete or change the base.
4. To insert a base in the consensus sequence, click between two bases, then insert the bases.

Note: The changed bases appear in lowercase.

Note: If the audit feature is enabled, you must enter a reason for each base change, base insertion, and base deletion.

Editing Sample Bases

You can edit sample bases in the same manner as consensus bases. However, only the sample whose base is edited and the consensus sequence are affected by the changes.

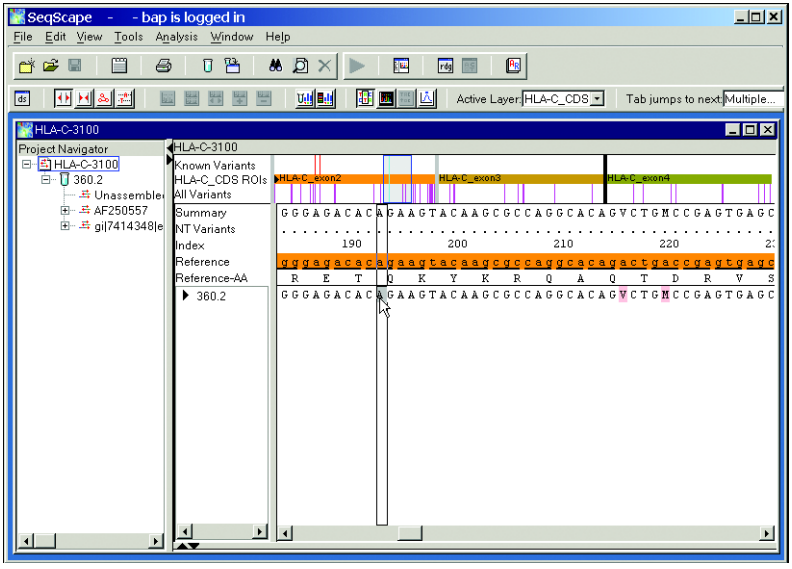
Editing a Consensus Sequence in the Project View

Note: Any changes are reflected in the sample sequences within the specimen and in the summary.

To insert, delete, or change a base in the Project view:

1. Select the project of interest in the navigation pane in the Project window.

   The project view opens displaying the consensus sequences for each specimen.



2. To change or delete a consensus base, click the base you want to edit, then delete or change the base.
3. To insert a base in the consensus sequence, click between two bases, then insert the base.

   Note: The changed bases appear in lowercase.

4. To delete a space, click the space to select, then press the **Delete** key.
5. To insert a space, click where you want to insert a space, then press the dash key or space bar.

Note: If the audit feature is enabled, you must enter a reason for each base change, base insertion, and base deletion.

Adjusting the Clear Range

About the Clear Range

Sample data usually has unreadable or otherwise unusable sequence located at the beginning and end of the data. Inclusion of this data causes errors in the alignments and erroneous variant detection.

The clear range is the area of continuous sequence that is the most error free. In SeqScape software, the clear range is set automatically for all samples during the analysis based on the Analysis Settings for that sample. You can modify the clear range on a per-sample basis.

IMPORTANT! If you do not select Use Reference Trimming in the Analysis Settings, you should manually set the clear range to remove any sample data that lies 5´ of the 5´end or 3´of the 3´end of the reference, if needed. Any sample data that is outside the reference is not aligned.

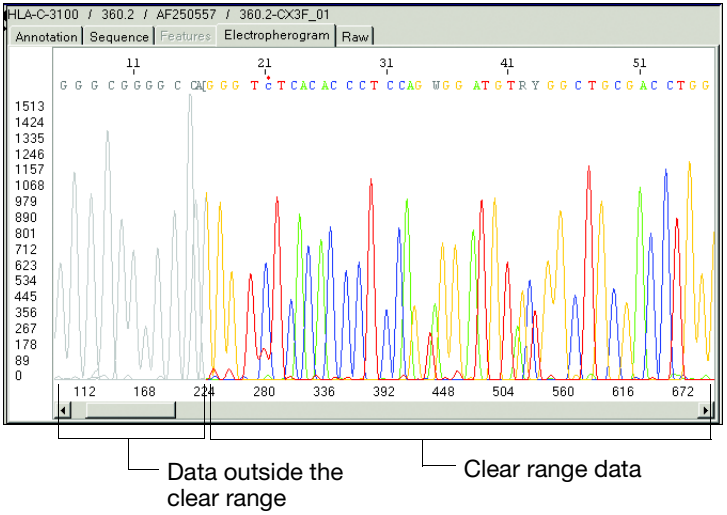


Figure 8-3 Clear Range Data

After changing the clear range, the specimen is automatically reassembled, then realigned and recompared to the reference.

The three methods for changing the clear range involve using the:

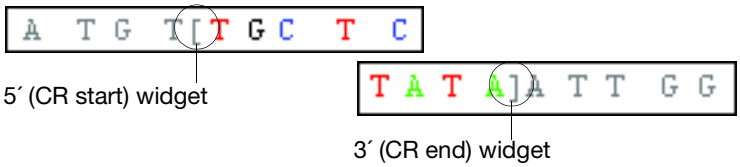
- Clear range widgets
- Mouse
- Set Clear Range dialog box

Using the Clear Range Widget

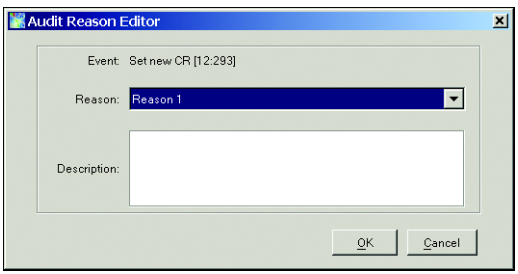
To use the Clear Range widget to adjust the clear range:

1. Open a sample from within a project.
2. Select the **Electropherogram** tab.
3. Locate and select the 5´ (CR start) or 3´(CR end) widget.

   The widget turns from gray to black, when selected.



4. Drag the widget along the bases to the right or left, as desired, then release the cursor.
5. If the audit feature is enabled, an Audit Reason Editor opens.



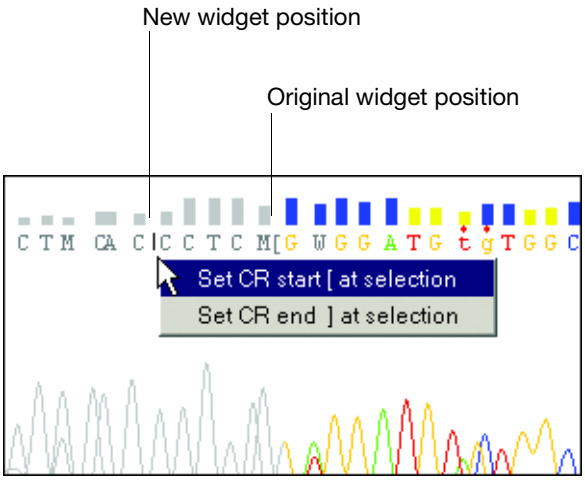
6. Complete the Audit Reason Editor dialog box, then click **OK**.

   The new clear range is displayed.
7. Repeat the process to define a new clear range for the opposite end.

Using the Mouse

To use the mouse to adjust the clear range:

1. Open a sample from within a project
2. Select the **Electropherogram** tab.
3. Right-click between two bases where you want to move the 5´ (CR start) or 3´(CR end) widget.



4. Do one of the following:

If you are moving the…	Then select…
CR start widget	**Set CR start [at selection**
CR end widget	**Set CR end] at selection**

5. If the audit feature is enabled, select a reason in the Audit Reason Editor, then click **OK**.

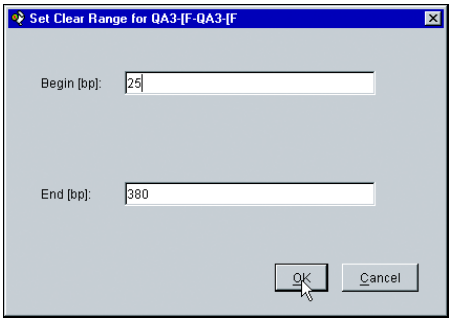
   The new clear range is displayed.

6. Repeat the process to define a new CR widget position for the opposite end.

Using the Set Clear Range Dialog Box

To use the dialog box to adjust the clear range:

1. Open a sample from within a project.
2. In the Electropherogram view or Specimen view, determine your new beginning and ending base numbers.
3. Select **Tools > Set Clear Range**.



4. Enter the values determined in step 2, then click **OK**.
5. If the audit feature is enabled, select a reason in the Audit Reason Editor, then click **OK**.

   The new clear range is displayed.

Editing Variants

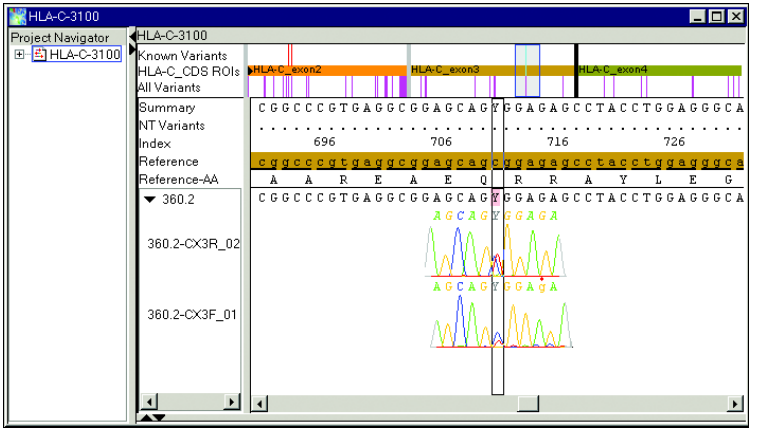
After you clean up errors in the sequences, you can view and edit the variants. Two methods to review variants follow.

Method 1

To view and edit variant data:

1. Open a project of interest.
2. Click a summary base.
3. Click the triangle that appears next to the specimen name to view the electropherogram snippets.
4. Edit bases or spaces in the specimen consensus sequences or the shown sample data.
5. Press **Tab** to move to the next variant or press **Shift-Tab** to move to the previous variant to view or edit more positions.

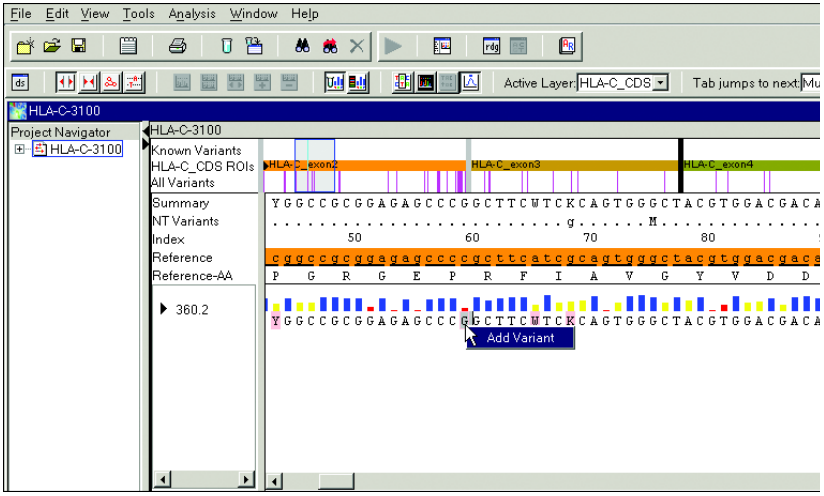
   Note: Pressing Ctrl+Z centers the selected column in the display, even if snippets are not showing.



Method 2

To view and edit variant data:

1. Open the project of interest.
2. Select **Analysis** > **Report Manager**.
3. In the navigation pane, select the report you want to view.
4. Select **Window** > **Tile**.
5. Review the positions by selecting a base change in the Mutations table. This adjusts the alignment view to the correct position in the alignment.
6. To add an unknown variant to the RDG, right-click the unknown variant position in a consensus sequence in the project alignment, then click **Add Variant**.



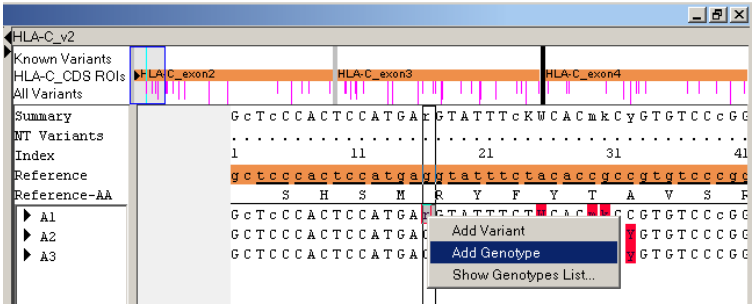
Saving Your Data

When you finish, save your project. Select **File** > **Save Project** or click (Save Project).

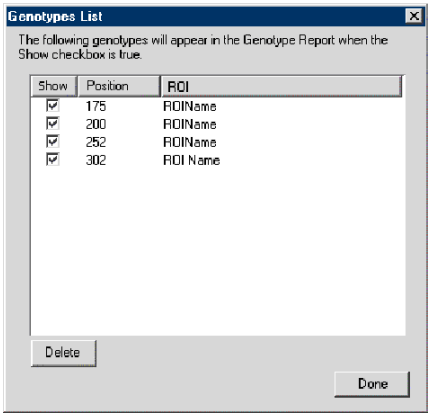
Adding a Genotype

To add a genotype to the Genotype Report:

1. In the Project View, select a specimen base, right-click, and then select **Add Genotype**.



2. To view the genotype list, select the specimen base, right-click, and then select **Show Genotypes List**.



3. Select the **Show** checkbox to display that position in the Genotyping Report. See “Genotyping Report” on page 7-39.

Exporting and Printing Data and Reports

9

This chapter contains:

Exporting Data Files 9-3
Exporting Reports 9-8
Printing Data and Reports 9-11

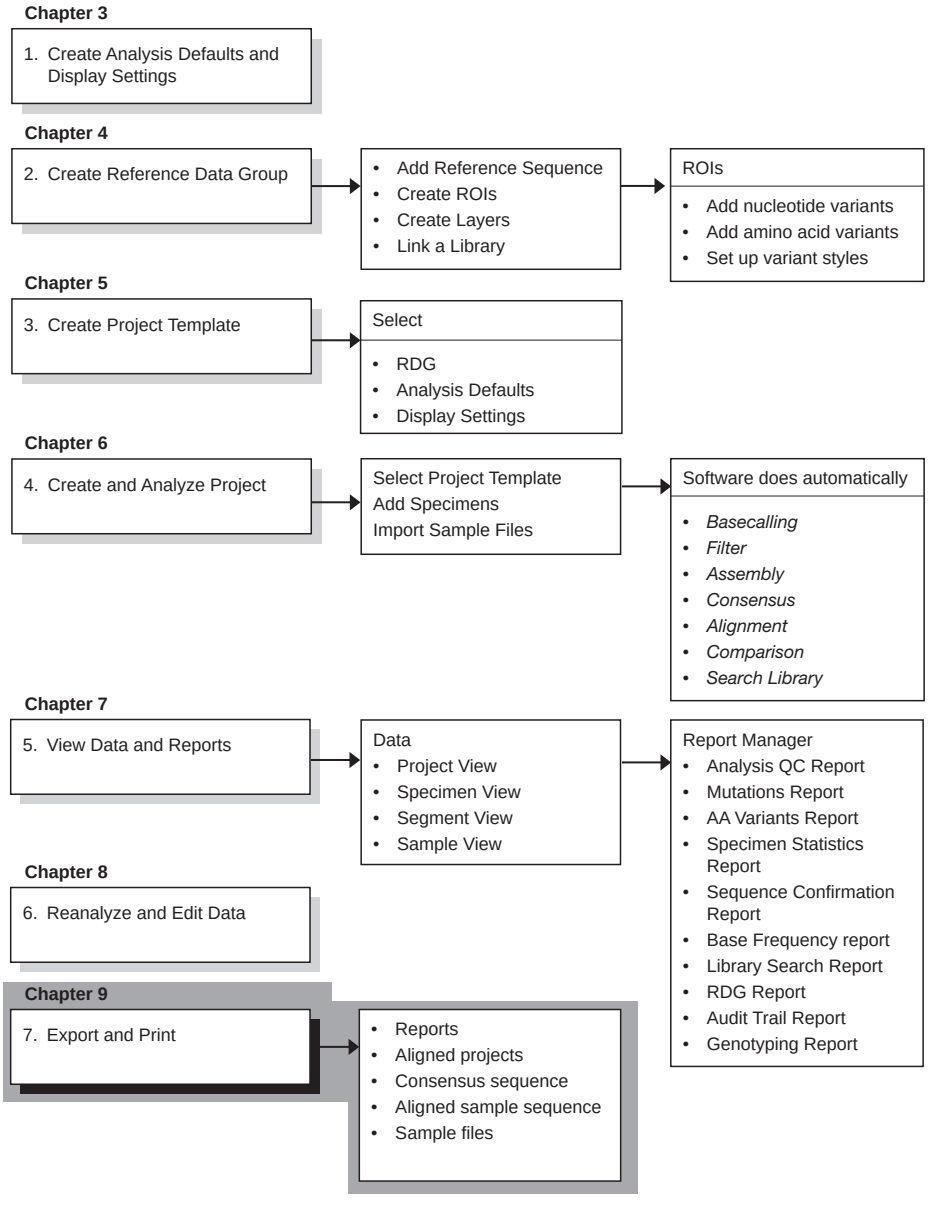


Figure 9-1 Showing the Printing and Exporting Steps

Exporting Data Files

File Names

The default file name uses the project name and the report type.

Do not use the following characters in any file name:
\ / : * ? “ < > | & and space

Format Options

You can export a project, specimen, segment, or sample file. Table 9-1 summarizes the available format options. Header and footer information is not incorporated in any data file.

Note: Only one data file can be exported at a time.

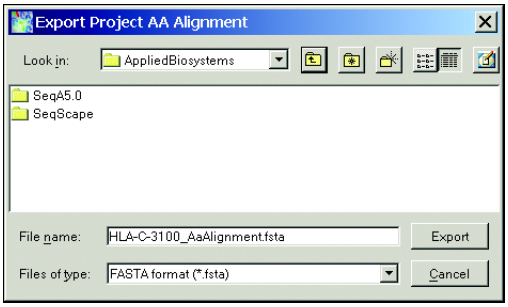
Table 9-1 Export and File Format Options

Export Option	File Format Options
Project	
Project Alignment-Nucleotides	FASTA
Project Alignment-Amino Acids	
Specimen and Segment	
Consensus Sequence	FASTA, SEQ, or QUAL
Aligned Sample Sequence	FASTA
Sample	
Sample Sequence File	FASTA, SEQ, AB1, or PHD

Exporting a Project Alignment

To export a project alignment:

1. Open the project of interest.
2. In the navigation pane, select the project icon.
3. Select **File > Export > Project Alignment-Nucleotides** or **Project Alignment-Amino Acid**.



4. Complete the Export Project dialog box:
   a. Select a folder location to store the project view.
   b. Change the file name, if desired.

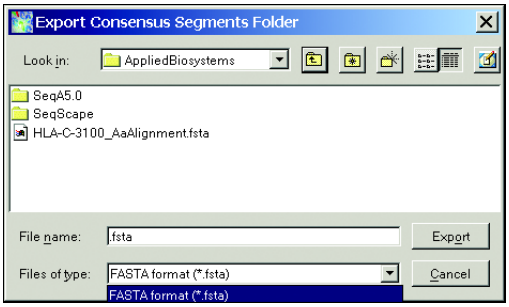
   Note: The default file name uses the project name with the element type suffix and the FASTA extension.

   c. Click **Export**.

Exporting a Specimen

To export a specimen:

1. Open the project of interest.
2. In the navigation pane, select a Specimen icon.
3. Select **File > Export > Consensus Sequence** or **Aligned Sample Sequence**.



4. Complete the Export Consensus dialog box:
   a. Select a folder location to store the file.

   Note: The default file name uses the project name with the element type suffix and the FASTA extension.

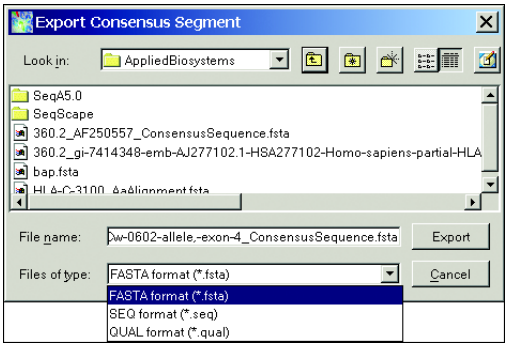
If the number of segments in a project is …	Then …
One	Change the file name, if desired.
Two or more	Do *not* type a file name. **Note:** The individual segment names are used. Any name you type is ignored.

   b. For the Consensus Sequence option, select a file format in the Files of type drop-down list.
   c. Click **Export**.

Exporting a Segment

To export a segment:

1. Open the project of interest.
2. In the navigation pane, select a Segment icon.
3. Select **File > Export > Consensus Sequence** or **Aligned Sample Sequence**.

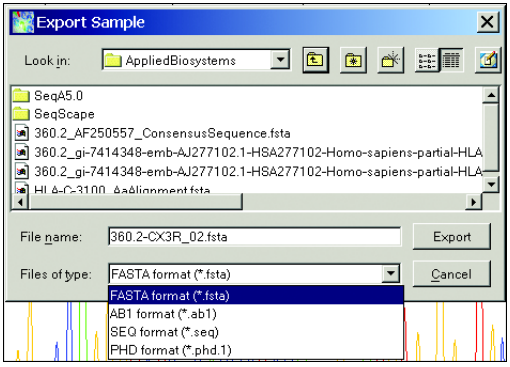


4. Complete the Export dialog box:
   a. Select a folder location to store the file.
   b. Change the file name, if desired. The default file name uses the segment name and the FASTA extension.
   c. For the Consensus Sequence option, select a file format in the Files of type drop-down list.
   d. Click **Export**.

Exporting a Sample

To export a sample:

1. Open the project of interest.
2. In the navigation pane, select a Sample icon.
3. Select **File > Export > Sample Sequence File**.



4. Complete the Export dialog box:
   a. Select a folder location to store the file.
   b. Change the file name, if desired. The default file name uses the sample name and the FASTA extension.
   c. Select a file format in the Files of type drop-down list.
   d. Click **Export**.

Exporting Reports

File Names

The default file name uses the project name and the report type.

Do not use the following characters in a file name:
\ /: * ? > | and space

Format Options

You can export generated reports as portable document format (pdf), text, HTML, or XML (Table 9-2).

Note: When selecting between HTML and XML, use HTML for standard display and XML for scripting applications.

Table 9-2 File Formats and Corresponding Application Options

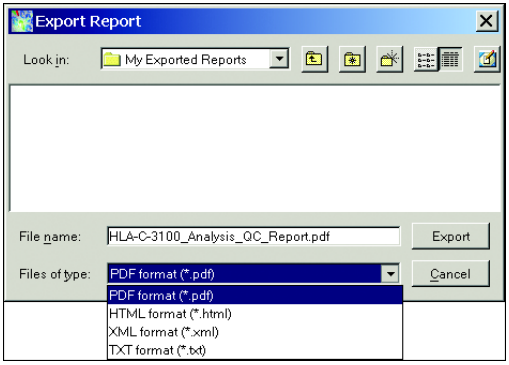
File Format	Open with…
PDF (default)	Adobe® Acrobat® Reader™
HTML*	A web browser or any software that is able to display HTML files
XML	A web browser
TXT	Notepad, Wordpad, Microsoft® Word, or any text-compatible software

*When exporting the report as HTML, a folder is automatically created that may contain more than one HTML file. The file that uses only the report name contains all the data from the report.

Exporting a Report

To export a report:

1. Open the project of interest, then click .
2. In the navigation pane, select a report type.
3. Customize the report, if desired. (See “Customizing the Reports” on page 7-42.)
4. Select **File** > **Export** > **Report**.



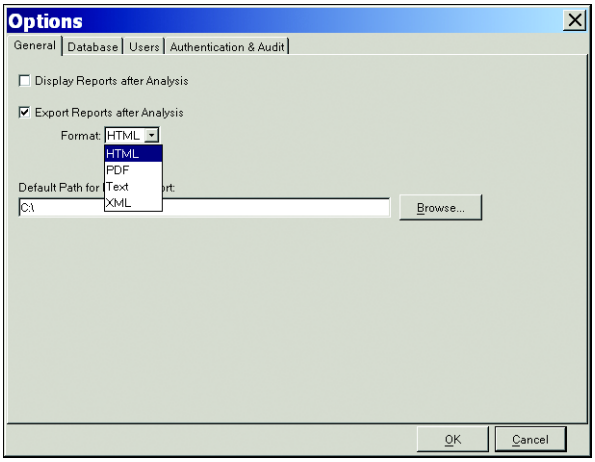
5. Complete the Export Report dialog box:
   a. Select a folder location to store the report.
   b. Change the file name of the report, if desired. The default file name uses the project name, the report type, and the .pdf extension.
   c. Select a file format in the Files of type drop-down list.
   d. Click **Export**.

Exporting All Reports Automatically

Reports can be automatically exported after analysis.

To set up for automatic exporting of reports:

1. Select **Tool** > **Options**.



2. Complete the General tab of the dialog box:
   a. Select the **Display Reports after Analysis** check box, if desired.
   b. Select the **Export Reports after Analysis** check box, then select an export format from the drop-down list.
   c. Define a default location to save the exported files.
   d. Click **OK**.

Printing Data and Reports

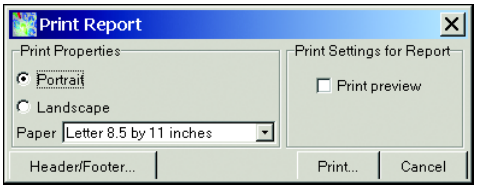
You can print any viewable screen in a WYSIWYG (what you see is what you get) manner within the SeqScape software on one of the recommended printers (HP 8100, 4500, 990cxi, and Epson 980). You can print project, specimen, segment, and sample views, as well as the reports for a project.

Customizing Header and Footer Display

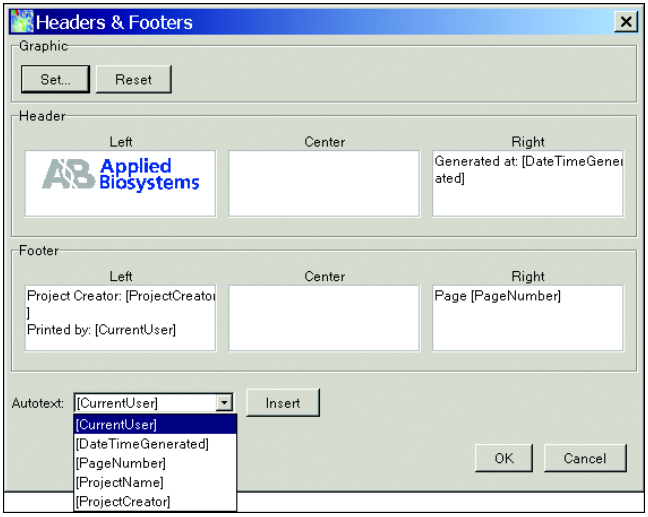
Default header and footer information is included in all exported and printed reports and in printed data views. However, headers and footers are not included in exported data files.

To customize the header/footer display in printed and exported reports:

1. Select **File** > **Print**.



2. Click **Header/Footer**.



3. To change the graphic, if desired:
   a. In the Graphic section, click **Set**.
   b. In the dialog box, locate then select a graphic file.
   c. Click **OK**.

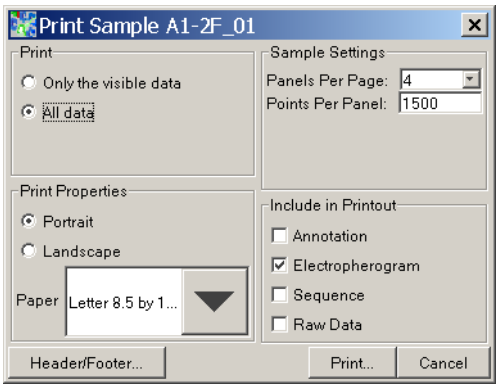
   Note: The graphic is displayed in the Headers & Footers dialog box and in the upper left corner of printed or exported reports.
4. To change the header and/or footer information, do one of the following:
   - Type text into any of the header and/or footer text boxes.
   - Use the autotext variables from the Autotext drop-down list. (Insert the cursor in a text box, select an autotext option in the drop-down list, then click **Insert**).
   - Use a combination of typing text and using the autotext variables.
5. Save the changes:
   a. Click **OK** to close the Header & Footer dialog box.
   b. Click **Print** in Print Report dialog box.
   c. In the Print dialog box, click:
      - **Cancel** to save the changes without printing
      - **OK** to save the changes and print

Printing Views of a Sample File

To print different views of a sample file:

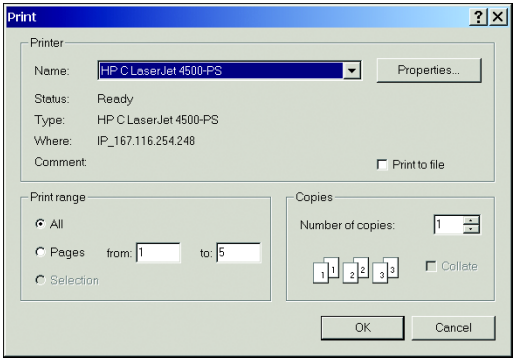
1. Open the project of interest.
2. In the navigation pane, select a sample to print.

3. Select **File > Print** or click .



4. Complete the dialog box:
   a. In the Print section, select **All data**.
   b. In the Print Properties section, select the paper orientation and size.
   c. In the Sample Settings section, select a value in the Panels Per Page drop-down list. The range is 1 to 4, and the default is 4.
   d. Select a value in the Points Per Panel value box. The range is 100 to 12000 and the default is 1500 (about 120 bases).
   e. In the Include a Printout section, select the views you want to print.
   f. If you are printing a Segment assembly, type a new value in the Bases per panel field in the Print Settings for Project section.

g. Click **Print**.



5. Select a printer, then click **OK**.

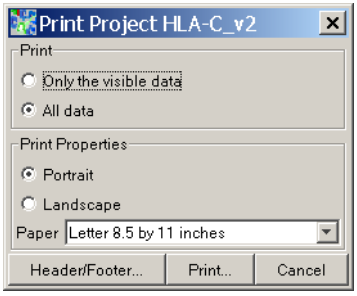
   Both print dialog boxes close and printing begins.

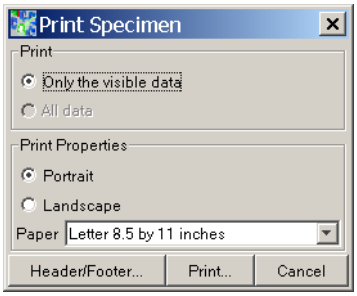
Printing Various Views of a Project

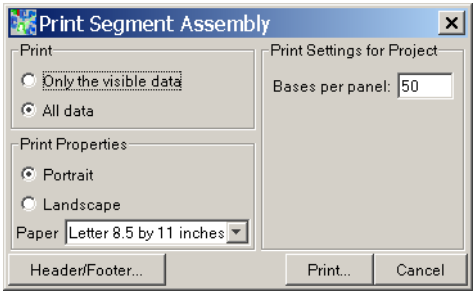
To print different views of a project:

1. Open the project of interest.
2. In the navigation pane, select a view (Project, Specimen, or Segment) to print.
3. If you are using WYSIWYG, scroll to the area of the view you want to print.

4. Select **File** > **Print** or click .

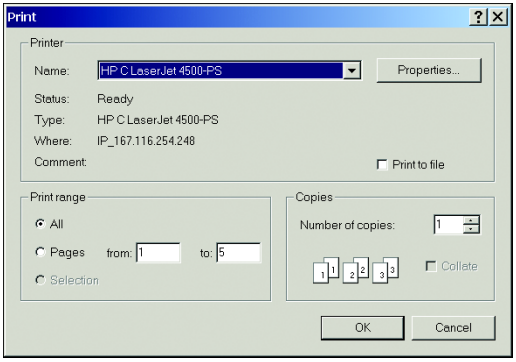






5. Complete the dialog box:
   a. In the Print section, select **Only the visible data** (WYSIWYG) or **All data** (if available).
   b. In the Print Properties section, select the paper orientation and size.
   c. If you are printing a Segment assembly, type a new value in the Bases per panel field in the Print Settings for Project section.

d. Click **Print**.



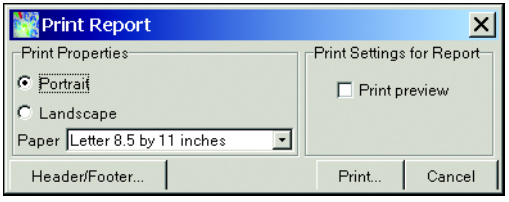
6. Select a printer, then click **OK**.

   Both print dialog boxes close and printing begins.

Printing a Report

To print a report:

1. Open the project of interest, then click .
2. In the navigation pane, select a report type.
3. Customize the report, if desired (see “Customizing the Reports” on page 7-42).
4. Select **File** > **Print**.

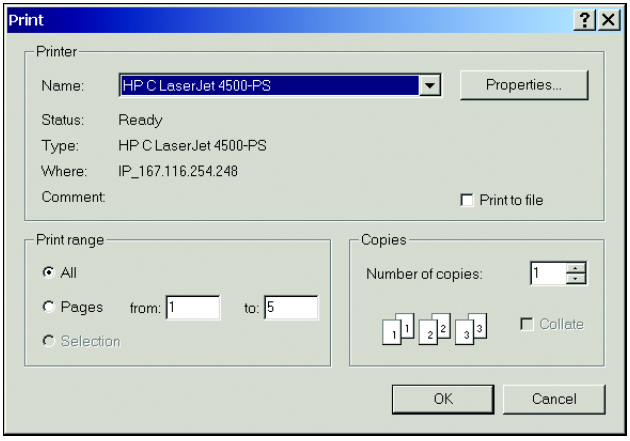


5. Complete the Print Report dialog box:
   a. In the Print Properties section, select the paper orientation and size.
   b. In the Print Settings for Report section, select **Print Preview**, if desired.
   c. Click **Print**.

6. Use the following table to determine your next step:

If the Print preview option was …	Proceed to step …
Not selected	7
Selected	8

7. The Print dialog box opens:



Select the printer and define the page range, then click **OK**.

8. The Report Preview dialog box opens. Use the command buttons as described in Table 9-3 on page 9-19.

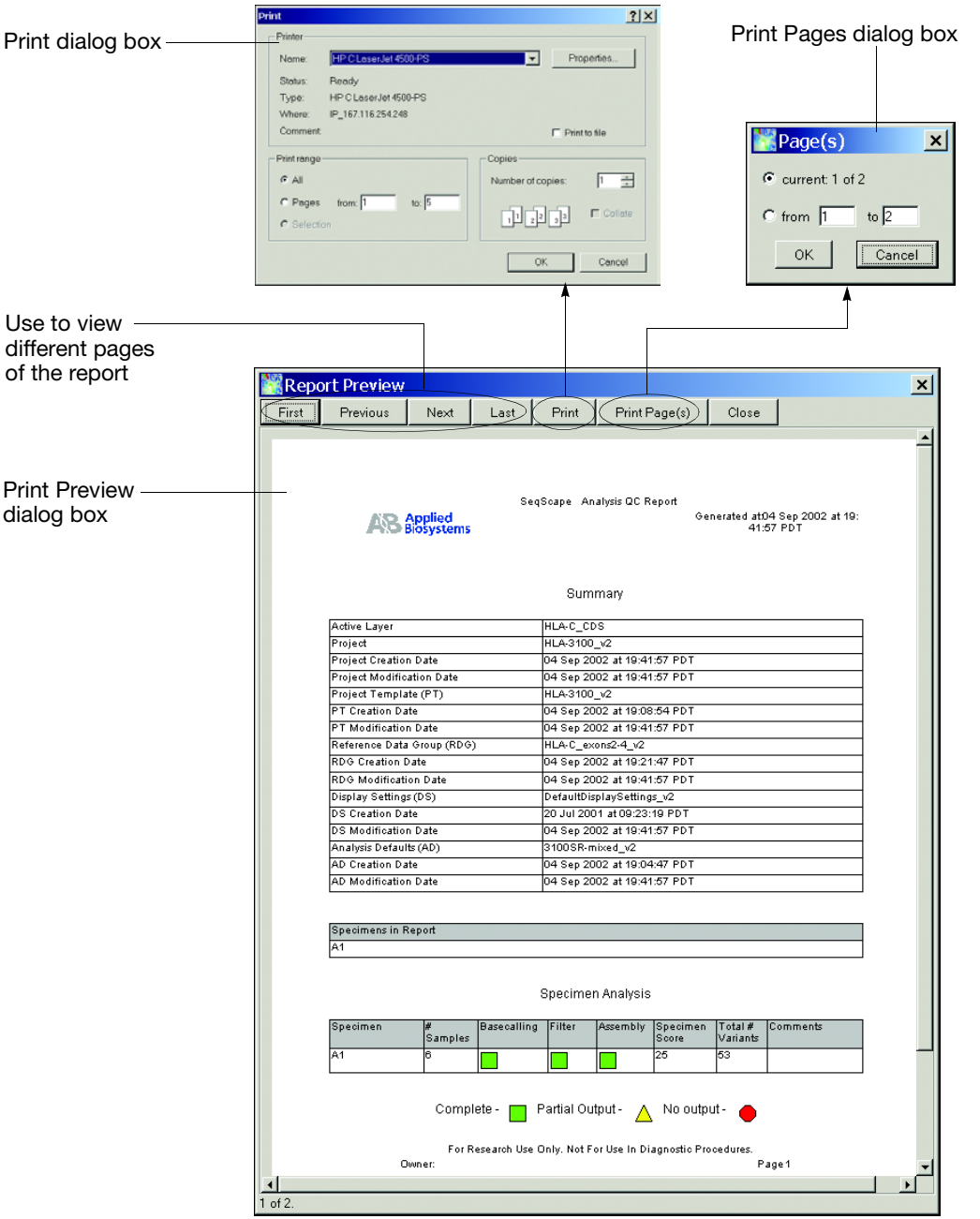


Table 9-3 Report Preview Button Functions

Button(s)	Function
First, Previous, Next, and Last	Displays the various pages in a report (only one page is visible at a time).
Print	Opens the Print dialog box. Select a printer, then click **OK** to print the report.
Print Pages	Opens the Page(s) dialog box. Set the page range, then click **OK**. In the Print dialog box, click **OK** to print the report. **Note:** Page(s) dialog box settings override the settings in the standard Print dialog box.
Close	Closes the preview window without printing the report.

Sample and Consensus Quality Values

10

This chapter contains:

Types of Quality Values (QVs) 10-2
Sample Quality Values 10-3
Consensus Quality Values 10-5
Displaying Quality Values. 10-6
Editing Bases with Quality Values 10-10
Cumulative Quality Value Scoring in Reports 10-11

Types of Quality Values (QVs)

Table 10-1 summarizes the types of QVs and where they are displayed.

Table 10-1 Quality Value Types

Quality Value Type	Definition	Location
Sample QV	A per-base estimate of basecaller accuracy.	• Sample view • Specimen view • Project view
Sample Score	The average quality value of the bases in the clear range sequence for that sample.	Specimen Statistics report
Consensus QV	A per-base estimate of the accuracy of the consensus-calling algorithm.	• Specimen view • Project view
Consensus Score	The average quality value of the bases in the consensus sequence for that specimen.	• Analysis QC report • Specimen Statistics report
Mutation QV	A per-base estimate of basecaller accuracy.	Mutations report
QV for deletion mutation	Average of the quality values for the bases to the left and right of the deletion.	Mutations report
FGR Avg QV	Average quality value of the consensus sequence across the target region	Genotyping report
5’ Reg Coverage Avg QV	Average quality value of the consensus sequence at the 5’ region	Genotyping report
Exon Avg QV	Average quality value of the consensus sequence at the exon region	Genotyping report
Intron Avg QV	Average quality value of the consensus sequence at the intron region	Genotyping report

Sample Quality Values

Sample Quality Values

A sample quality value (SQV) is a per-base estimate of the basecaller accuracy. There are two types of basecallers that generate SQVs:

- KB – A new algorithm that identifies mixed or pure bases, and generates sample quality values.
- ABI – Algorithm used in ABI PRISM® Sequencing Analysis Software v3.7 that identifies pure bases. Then the TraceTuner™ software identifies mixed bases and generates sample quality values.

KB and ABI algorithms can produce slightly different SQVs.

Interpreting the Sample Quality Values

Per-base SQVs are calibrated on a scale corresponding to:

$$QV = -10\log_{10}(Pe)$$

Where *Pe* is the probability of error of the basecall.

The KB basecaller produces a QV range of 1 to 99, with 1 being low confidence and 99 being high confidence. See Table 10-2, “Quality Values and Probabilities of Error,” on page 10-4 for the probability of basecall errors for QVs ranging from 1 to 99.

Mixed base calls yield lower SQVs than pure base calls.

The typical QV range for good data is 25 to 50.

The size and color of QV bars for QV 50 to 99 are identical. To view the actual QV number, see “Displaying the Quality Bars and Values” on page 10-9.

Sample Score

A sample score is generated from SQVs. It is the average quality value of the bases in the clear range sequence for a sample.

Table 10-2 Quality Values and Probabilities of Error

Quality Value = $-10\log_{10}(Pe)$					
where *Pe* is probability of error					
KB basecaller generates QVs from 1 to 99					
Typical high quality pure bases will have QV 20- 50					
Typical high quality mixed bases will have QV 10-50					
Size and color of QVs bars are identical for QVs 50-99					
QV	***Pe***	**QV**	***Pe***	**QV**	***Pe***
1	79%	21	0.790%	41	0.0079%
2	63%	22	0.630%	42	0.0063%
3	50%	23	0.500%	43	0.0050%
4	39%	24	0.390%	44	0.0039%
5	31%	25	0.310%	45	0.0031%
6	25%	26	0.250%	46	0.0025%
7	20%	27	0.200%	47	0.0020%
8	15%	28	0.150%	48	0.0015%
9	12%	29	0.120%	49	0.0012%
10	10%	30	0.100%	50	0.0010%
11	7.9%	31	0.079%	60	0.0001%
12	6.3%	32	0.063%	70	0.00001%
13	5.0%	33	0.050%	80	0.000001%
14	4.0%	34	0.040%	90	0.0000001%
15	3.2%	35	0.032%	99	0.00000012%
16	2.5%	36	0.025%		
17	2.0%	37	0.020%		
18	1.6%	38	0.016%		
19	1.3%	39	0.013%		
20	1.0%	40	0.010%		

Consensus Quality Values

A consensus quality value (QV) is a per-base estimate of the accuracy of the consensus-calling algorithm. If the SQVs are generated from the KB basecaller, then the KB consensus-calling algorithm is used to generate the QVs. If the SQVs are generated from an ABI basecaller and TraceTuner, then the TraceTuner consensus-calling algorithm is used to generate the QVs.

The KB and TraceTuner consensus-calling algorithms can produce slightly different consensus QVs.

Interpreting the Consensus Quality Values

The degree of certainty of either consensus-calling algorithm is reflected by the per-base consensus QVs. A consensus QV is derived from a number of factors:

- How large a quality-value discrepancy exists between calls from the individual sample sequence strands
- The possible redundancy of calls from strands in the same orientation
- The possibility that the basecaller missed a mixed base

The possible values for the QVs range from 1 to 50. Higher numbers indicate calls that the algorithm determined with a measure of confidence, while lower numbers indicate calls that might require user inspection to verify the correct answer. The consensus quality values are roughly calibrated to follow the same scale as the per-base sample quality values.

Consensus Score

A consensus score is generated from consensus QVs. It represents the average quality value of the bases in the consensus sequence for a specimen.

Displaying Quality Values

QVs are displayed as bars above each base in a sample (Figures 10-1 and 10-2). The height and color of a bar indicates its value. The taller the bar, the higher the QV. The color of a bar, which is associated with its value, is editable in the Display Settings.

Note: QV bar height and color are identical for QVs 50 to 99.

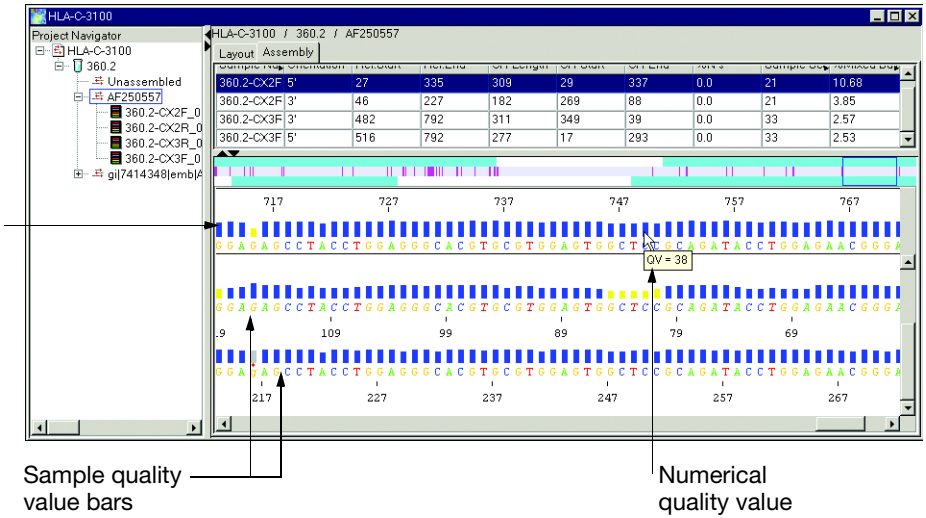


Figure 10-1 Example of QV Bars in the Specimen View

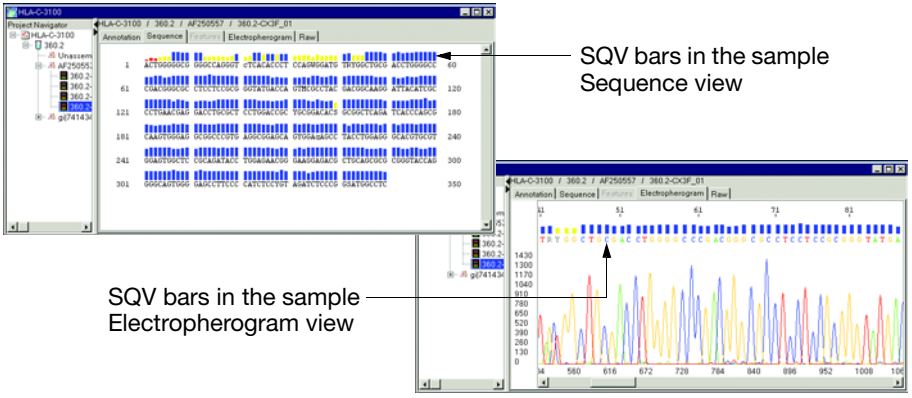


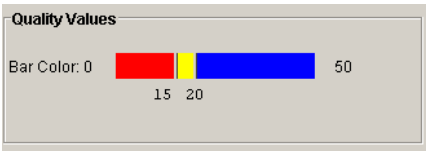
Figure 10-2 Examples of SQVs in the Sample View

Customizing the Quality Value Display

You can modify the low, medium, and high ranges and the color associated with a QV.

To modify the QV display:

1. Select **Analysis** > **Display Settings** or click .
2. Select the **Bases** tab.
3. In the Quality Values section, place the pointer between two colors (it becomes a double-headed arrow), then click the slider on the color bar and drag it to left or right to the desired value.

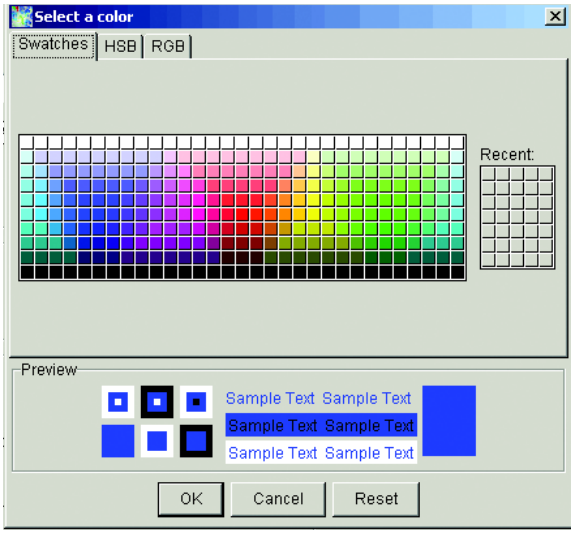


Use the criteria in the table below to define what values represent low, medium, and high ranges for your project.

QV Bar	Default Color and Range	Set the range to identify data that is …
Low	Red 0 to 14	Not acceptable
Medium	Yellow 15 to 19	Needs manual review
High	Blue 20 or higher	Acceptable

4. Change the colors that represent low, medium, and/or high QVs, if desired:
   a. Select the color in the Bar Code you want to change.

      The Select a color dialog box opens.



   b. Select a new color in the Swatches tab, or use the HSB or RGB tabs to define a new color.
   c. Click **OK**. The color dialog box closes.
5. Do one of the following:
   - Click **OK** to save the changes to the samples you are working with.
   - Click **Save to Manager As** to save the changes to the SeqScape Manager.

Displaying the Quality Bars and Values

If you do not see the QV bars when viewing samples or a consensus in a project, then follow the procedures below to display QV bars and values.

To view quality bars and values:

1. Open a project, then open a specimen of interest.
2. Select the segment of interest, then select the **Assembly** tab.
3. To view sample QVs, select **View > Show/Hide sample QV** or click .
4. To view consensus QVs, select **View > Show/Hide consensus QV** or click .
5. To obtain a numerical value for a particular bar, place the cursor over the bar for 2 sec. The value is automatically displayed.

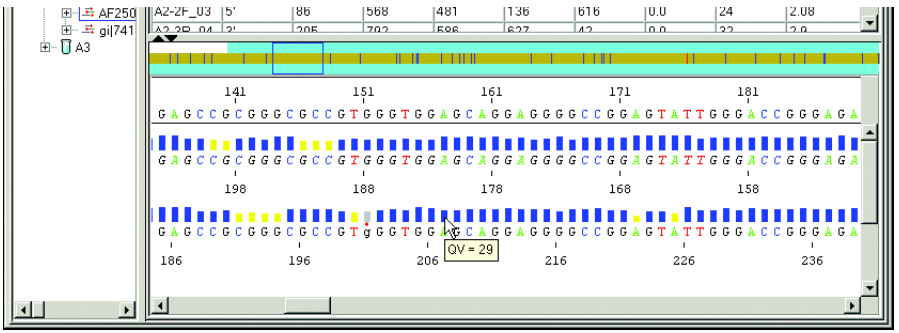


Figure 10-3 Displaying the Value of a Sample QV Bar

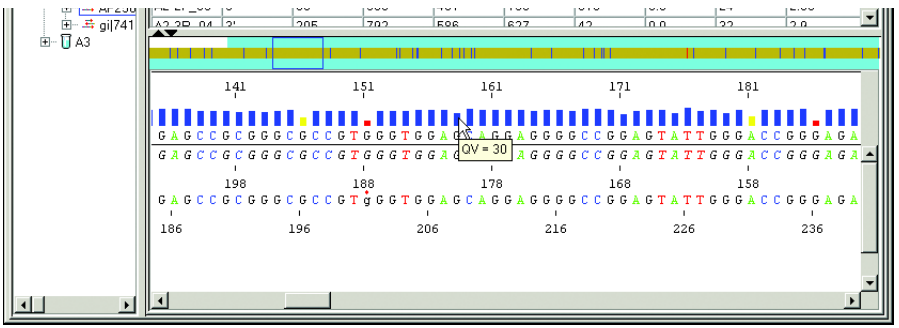


Figure 10-4 Displaying the Value of a Consensus QV Bar

Editing Bases with Quality Values

Changing, deleting, and inserting a base affect the consensus or sample QVs displayed.

Table 10-3 Results of Editing Bases with Quality Values

If …	Then …
The consensus-caller calls a base not present in all the samples	The new base is in uppercase in the consensus sequence and in lowercase in the samples that did not contain that basecall with a red dot.
You change a base	The new base is in lowercase and the SQV has the same value but is displayed as a gray bar.
You change a base back to the original call	The base appears in uppercase and the quality value bar color is restored.
You insert a base	The inserted base appears in lowercase and it has no SQV.
You delete a base	The quality value for the base disappears.
You reinsert a deleted base	The reinserted base appears in lowercase and it has no SQV.

Cumulative Quality Value Scoring in Reports

Quality values and scores are also displayed in several reports. To view the reports, select **Analysis > Report Manager** or click .

Analysis QC Report

Consensus scores in an Analysis QC report are shown as an average quality value across the consensus sequence for each specimen.

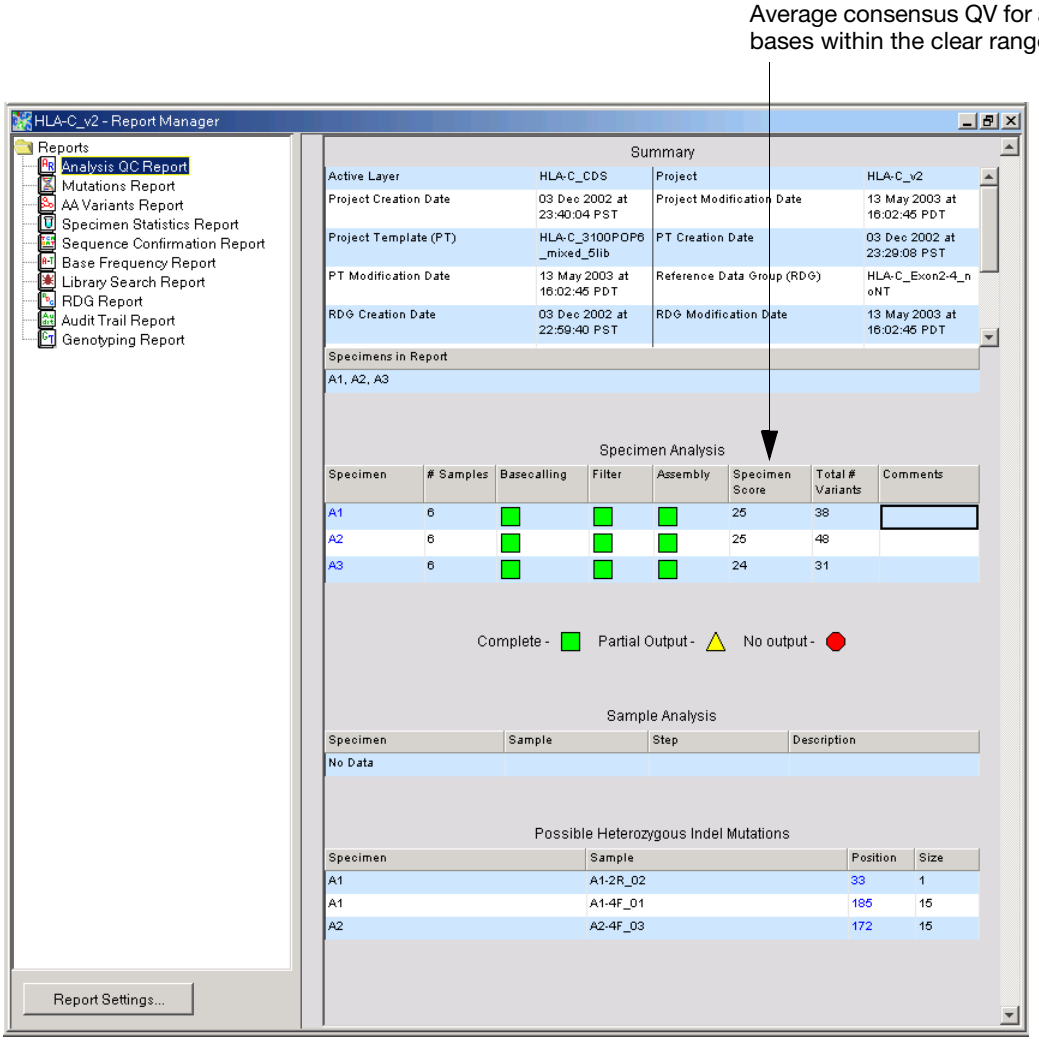


Figure 10-5 Analysis QC Report

Mutations Report

QVs for each mutation, and the average QV for the bases to the left and right of the deletion are provided in Mutations report.

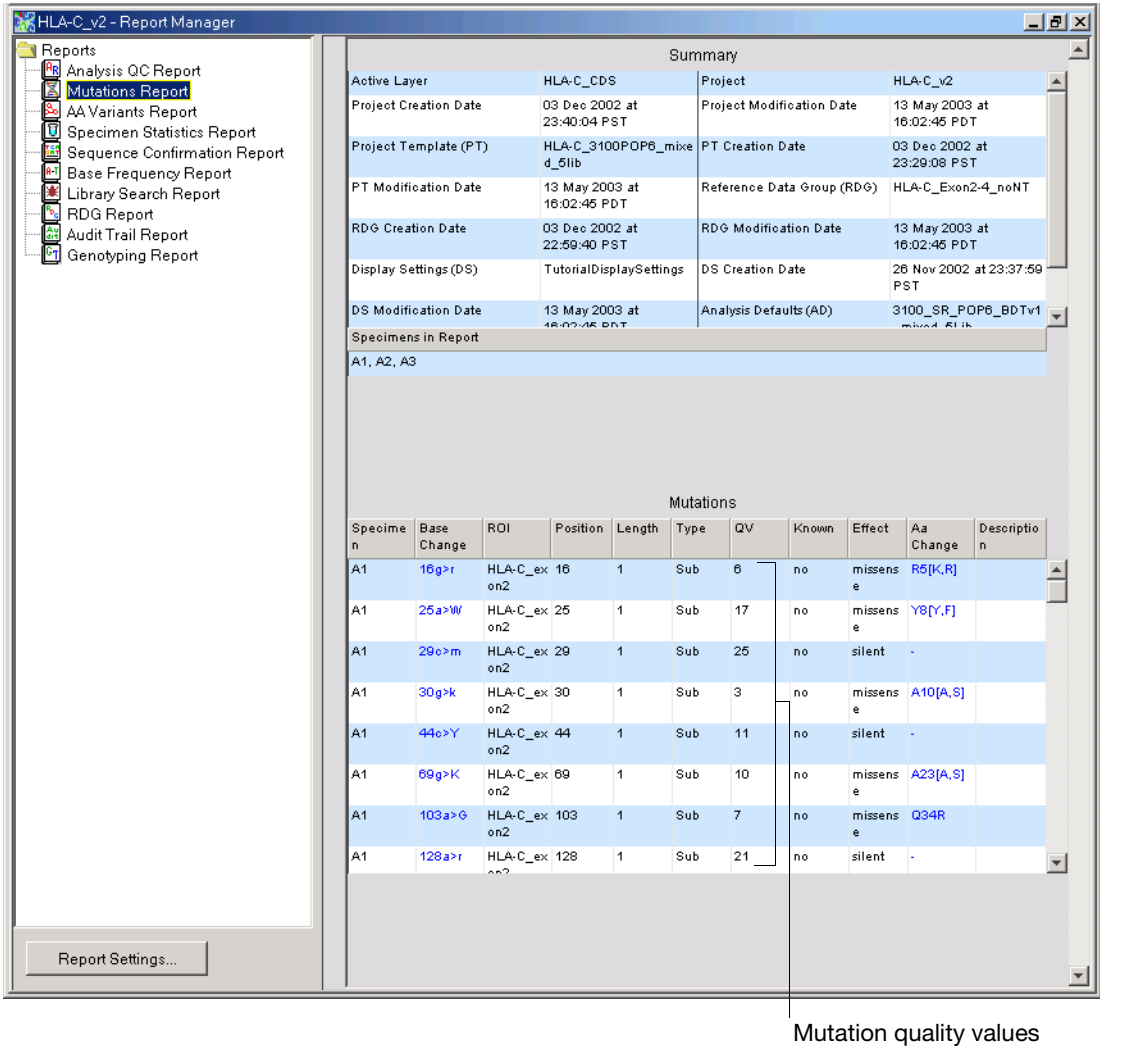


Figure 10-6 Mutations Report

Specimen Statistics Report

The Specimen Statistics table of this report displays the average consensus QV score for a segment in the Segment Score column.

The Sample Results table displays the average sample QV for the bases in the clear range in the Sample Score column.

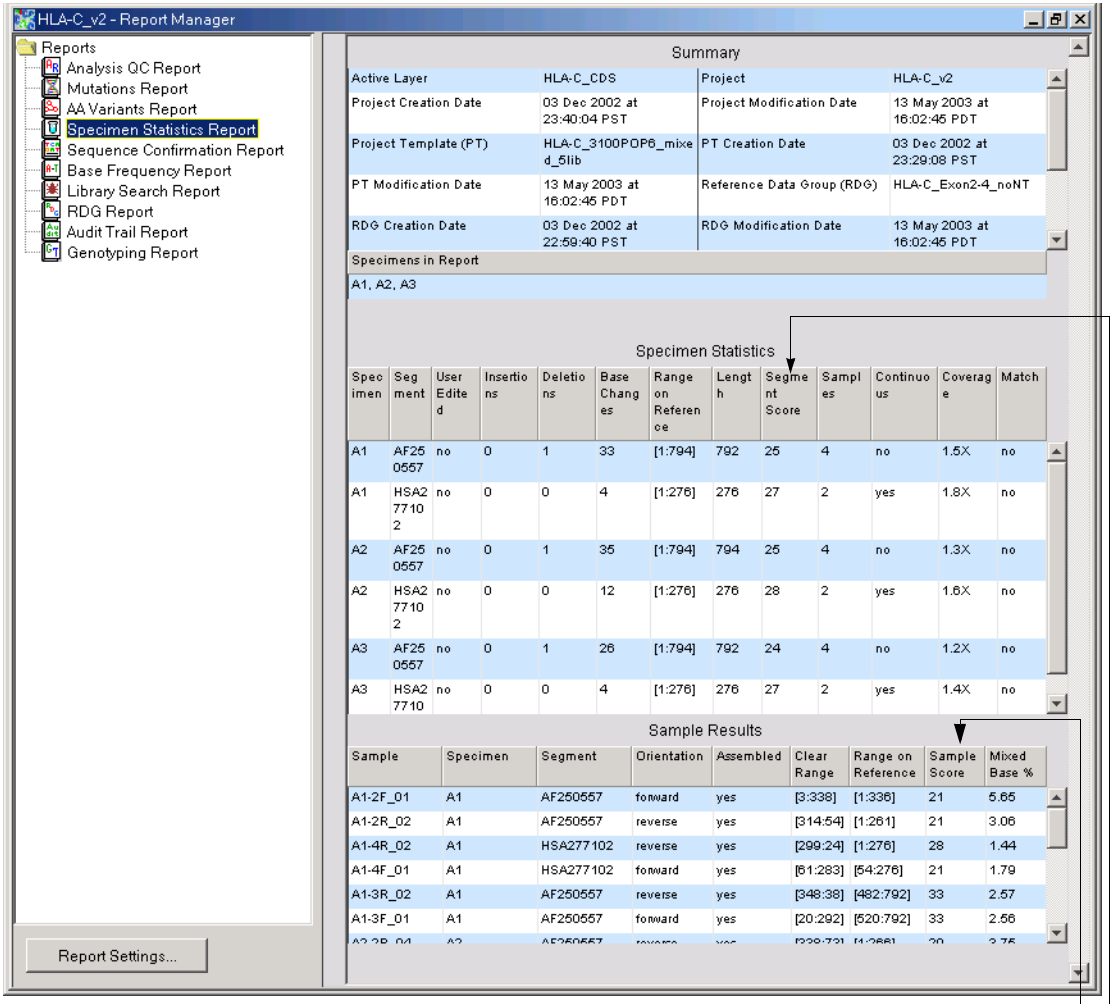

HLA-C_v2 - Report Manager
Reports
Analysis QC Report
Mutations Report
AA Variants Report
Specimen Statistics Report
Sequence Confirmation Report
Base Frequency Report
Library Search Report
RDG Report
Audit Trail Report
Genotyping Report
Report Settings...
Summary
Active Layer
HLA-C_CDS
Project
HLA-C_v2
Project Creation Date
03 Dec 2002 at 23:40:04 PST
Project Modification Date
13 May 2003 at 16:02:45 PDT
Project Template (PT)
HLA-C_3100POP6_mixed_5lib
PT Creation Date
03 Dec 2002 at 23:29:08 PST
PT Modification Date
13 May 2003 at 16:02:45 PDT
Reference Data Group (RDG)
HLA-C_Exon2-4_noNT
RDG Creation Date
03 Dec 2002 at 22:59:40 PST
RDG Modification Date
13 May 2003 at 16:02:45 PDT
Specimens in Report
A1, A2, A3
Specimen Statistics
Specimen
Segment
User Edited
Insertions
Deletions
Base Changes
Range on Reference
Length
Segment Score
Samples
Continuous
Coverage
Match
A1 AF250557 no 0 1 33 [1:794] 792 25 4 no 1.5X no
A1 HSA277102 no 0 0 4 [1:276] 276 27 2 yes 1.8X no
A2 AF250557 no 0 1 35 [1:794] 794 25 4 no 1.3X no
A2 HSA277102 no 0 0 12 [1:276] 276 28 2 yes 1.6X no
A3 AF250557 no 0 1 26 [1:794] 792 24 4 no 1.2X no
A3 HSA277102 no 0 0 4 [1:276] 276 27 2 yes 1.4X no
Sample Results
Sample
Specimen
Segment
Orientation
Assembled
Clear Range
Range on Reference
Sample Score
Mixed Base %
A1-2F_01 A1 AF250557 forward yes [3:338] [1:336] 21 5.65
A1-2R_02 A1 AF250557 reverse yes [314:54] [1:261] 21 3.06
A1-4R_02 A1 HSA277102 reverse yes [299:24] [1:276] 28 1.44
A1-4F_01 A1 HSA277102 forward yes [61:283] [54:276] 21 1.79
A1-3R_02 A1 AF250557 reverse yes [348:38] [482:792] 33 2.57
A1-3F_01 A1 AF250557 forward yes [20:292] [520:792] 33 2.56
Average sample QV for the bases is in the clear range
Average consensus QV score for a segment

Figure 10-7 Specimen Statistics Report

Genotyping Report

The Resequencing Coverage table lists several QVs.

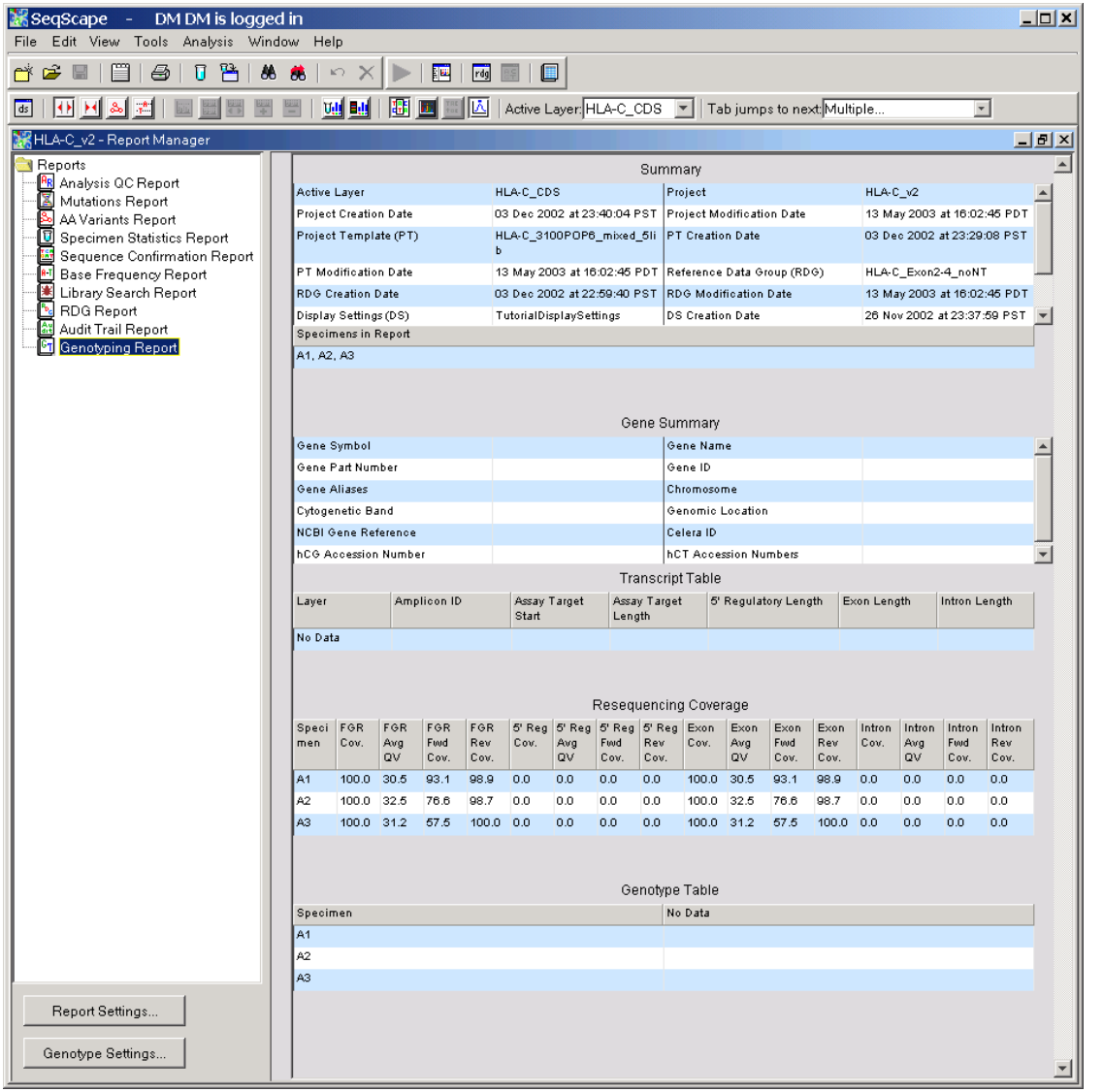

SeqScape - DM DM is logged in
File Edit View Tools Analysis Window Help
Active Layer: HLA-C_CDS
Tab jumps to next: Multiple...
HLA-C_v2 - Report Manager
Reports
Analysis QC Report
Mutations Report
AA Variants Report
Specimen Statistics Report
Sequence Confirmation Report
Base Frequency Report
Library Search Report
RDG Report
Audit Trail Report
Genotyping Report
Report Settings...
Genotype Settings...
Summary
Active Layer HLA-C_CDS
Project HLA-C_v2
Project Creation Date 03 Dec 2002 at 23:40:04 PST
Project Modification Date 13 May 2003 at 16:02:45 PDT
Project Template (PT) HLA-C_3100POP6_mixed_5lib
PT Creation Date 03 Dec 2002 at 23:29:08 PST
PT Modification Date 13 May 2003 at 16:02:45 PDT
Reference Data Group (RDG) HLA-C_Exon2-4_noNT
RDG Creation Date 03 Dec 2002 at 22:59:40 PST
RDG Modification Date 13 May 2003 at 16:02:45 PDT
Display Settings (DS) TutorialDisplaySettings
DS Creation Date 26 Nov 2002 at 23:37:59 PST
Specimens in Report
A1, A2, A3
Gene Summary
Gene Symbol
Gene Name
Gene Part Number
Gene ID
Gene Aliases
Chromosome
Cytogenetic Band
Genomic Location
NCBI Gene Reference
Celera ID
hCG Accession Number
hCT Accession Numbers
Transcript Table
Layer
Amplicon ID
Assay Target Start
Assay Target Length
5' Regulatory Length
Exon Length
Intron Length
No Data
Resequencing Coverage
Specimen FGR Cov. FGR Avg QV FGR Fwd Cov. FGR Rev Cov. 5' Reg Cov. 5' Reg Avg QV 5' Reg Fwd Cov. 5' Reg Rev Cov. Exon Cov. Exon Avg QV Exon Fwd Cov. Exon Rev Cov. Intron Cov. Intron Avg QV Intron Fwd Cov. Intron Rev Cov.
A1 100.0 30.5 93.1 98.9 0.0 0.0 0.0 0.0 100.0 30.5 93.1 98.9 0.0 0.0 0.0 0.0
A2 100.0 32.5 76.6 98.7 0.0 0.0 0.0 0.0 100.0 32.5 76.6 98.7 0.0 0.0 0.0 0.0
A3 100.0 31.2 57.5 100.0 0.0 0.0 0.0 0.0 100.0 31.2 57.5 100.0 0.0 0.0 0.0 0.0
Genotype Table
Specimen
No Data
A1
A2
A3

Figure 10-8 Genotyping Report

Automating Analysis 11

This chapter contains:

Integrating SeqScape and Data Collection Software 11-2
Before You Start 11-4
Required Files Created in SeqScape Software 11-6
Launching Autoanalysis Manger 11-7
Creating Required Files in the Data Collection Software 11-8
Scheduling and Starting a Run 11-22
Understanding the Autoanalysis Manager 11-24
File Sharing Between Data Collection and SeqScape Software .11-28

Integrating SeqScape and Data Collection Software

Overview

Sequencing data that is generated on Applied Biosystems 3730/3730*xl* DNA Analyzers and ABI PRISM® 3100/3100-*Avant* Genetic Analyzers can be automatically analyzed for use in the ABI PRISM® SeqScape® Software Version 2.1. Autoanalysis can be performed only on the same instrument computer that collected the sample files. You can configure the software packages to perform data collection and then data analysis without requiring user interaction.

Autoanalysis requires three software packages:

- **3730/3730*xl* or 3100/3100-*Avant* Data Collection software v2.0**

  Data collection software is used to run the instrument and collect fluorescent data from samples. For autoanalysis to occur, the data collection software must be set up properly to allow communication with the analysis software.

  Data collection software uses a data service. Data used for data collection as well as that created in SeqScape software can be accessed through the data service in data collection software.

- **Autoanalysis Manager**

  Autoanalysis Manager software is part of the integration between data collection, SeqScape, and GeneMapper™ software. It queues messages and tracks the status of their processing. Each message is considered a batch job, whether it contains a single sample, samples from a result group, or an entire run of samples.

  Autoanalysis Manager is installed by SeqScape or GeneMapper software when loaded on a system with data collection software.

- **A version of SeqScape software with no user interface**

  This version of SeqScape software is identical to the regular version of the software, except that no user interface exists. The Autoanalysis Manager opens and uses this version of software to analyze the data in the projects.

  The automated processing version and the standard version of SeqScape software are installed from the SeqScape Software installation CD.

IMPORTANT! **The data collection software must be running** when installing SeqScape software on a computer that is connected to a 3730/3730*xl* DNA Analyzer or 3100/3100-*Avant* Genetic Analyzer. If data collection software is not running, SeqScape software does not register with the Data Service. See Chapter 2, "Getting Started," for information on properly installing the software.

IMPORTANT! After the initial installation of SeqScape software, you must open Autoanalysis Manager software.

Software Relationships

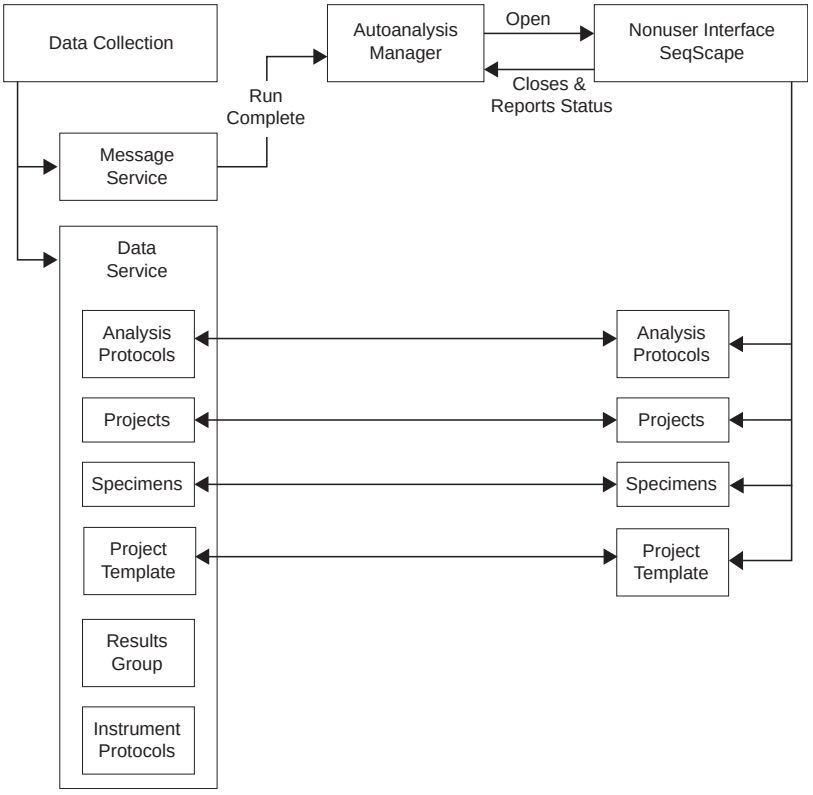


Figure 11-1 Relationships Between Seqscape and Data Collection Software

Before You Start

Successful automatic analysis requires that:

- SeqScape software is installed properly
- SeqScape software is registered and the appropriate user IDs have been created
- Autoanalysis Manager software is running
- The 3730/3730*xl* or 3100/3100-*Avant* instrument is set up to run, and samples are prepared

  For more information on setting up and using the 3730 Data Collection software, refer to the *Applied Biosystems 3730/3730xl Getting Started Guide* (PN 4347118).

  For more information on setting up and using the 3100 Data Collection software, refer to the *ABI PRISM 3100/3100-Avant User Guide* (PN 4347102).
- Files for a data collection software plate record are available

  For data collection and autoanalysis to be successful, each run of samples must have an instrument protocol, an analysis protocol, and a results group assigned within a plate record.

  The table below describes what each file specifies in the logical order of its use.

Table 11-1 File Specifications

File	Description	Created in
Instrument Protocol	Contains everything needed to run the instrument.	Data collection software
Analysis Protocol	Contains everything needed to analyze sequencing data.	Data collection software or SeqScape software
Results Group	Defines the file type, the file name, file save locations, default analysis protocols linked to sample injections, and user name and password.	Data collection software
Project Template	Contains RDG, analysis defaults, and display settings	SeqScape software

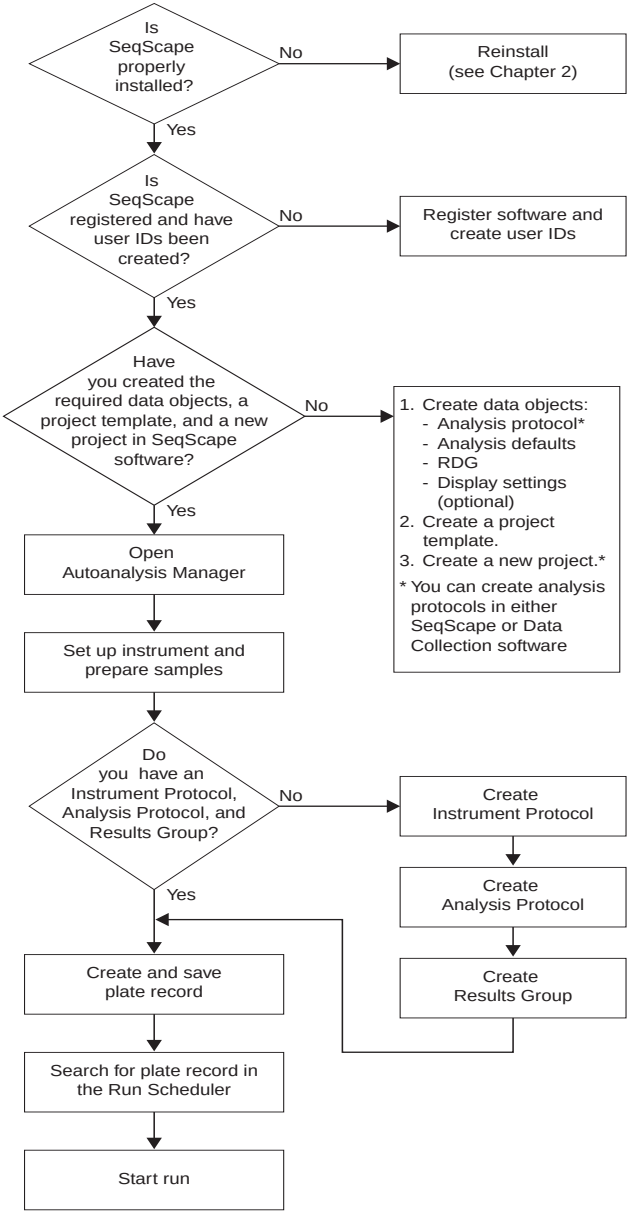

Is SeqScape properly installed?
No
Reinstall (see Chapter 2)
Yes
Is SeqScape registered and have user IDs been created?
No
Register software and create user IDs
Yes
Have you created the required data objects, a project template, and a new project in SeqScape software?
No
1. Create data objects:
- Analysis protocol*
- Analysis defaults
- RDG
- Display settings (optional)
2. Create a project template.
3. Create a new project.*
* You can create analysis protocols in either SeqScape or Data Collection software
Yes
Open Autoanalysis Manager
Set up instrument and prepare samples
Do you have an Instrument Protocol, Analysis Protocol, and Results Group?
No
Create Instrument Protocol
Create Analysis Protocol
Create Results Group
Yes
Create and save plate record
Search for plate record in the Run Scheduler
Start run

Figure 11-2 Autoanalysis Workflow

Required Files Created in SeqScape Software

During autoanalysis setup in the data collection software, you will be required to select data objects that have already been created. For example, if you choose to create a new project in the plate editor, you can select only a project template that is already registered in SeqScape software. You cannot create a project template in the plate editor.

Before continuing, we recommend that you:

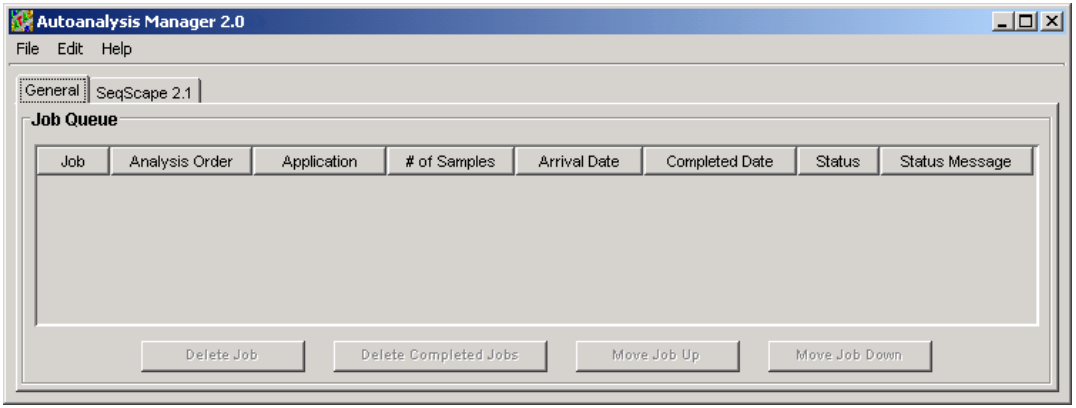
- Create the required data objects in SeqScape software:
  - Create an analysis protocol. An analysis protocol can be created in SeqScape or data collection software.
  - Create the analysis defaults and link them to an analysis protocol.
  - Create a Reference Data Group (RDG) based on your samples.
  - Customize the display settings or use the default settings.
- Create a project template and associate the objects you just created to it.
- Create a new project associated with the project template you just created.

Launching Autoanalysis Manger

Launch the Autoanalysis manager before continuing.

To Launch the Autoanalysis Manager:

1. Select **Start** > **Applied Biosystems** > **Autoanalysis Manager** > **Autoanalysis Manager 2.0**.



IMPORTANT! Do not run other applications when the Autoanalysis manager and the data collection software are running. Never open a project while it is in use by the Autoanalysis manager.

Creating Required Files in the Data Collection Software

For More Information

For more information on setting up and using the 3730/3730*xl* DNA Analyzer and the 3730 Data Collection software, refer to the *Applied Biosystems 3730/3730xl Getting Started Guide* (PN 4347118).

For more information on setting up and using the 3100/3100-*Avant* Genetic Analyzer and the 3100 Data Collection software, refer to the *ABI PRISM 3100/3100-Avant User Guide* (PN 4347102).

If the Files Already Exist

If the appropriate instrument protocol, analysis protocol, and results group have been created, proceed to "Creating a Plate Record" on page 11-18.

Creating an Instrument Protocol

Note: The figures used in this section are from the 3730 data collection software v2.0.

To create an instrument protocol:

1. In the navigation pane, click the **Protocol Manager** icon to display the Protocol Manager.

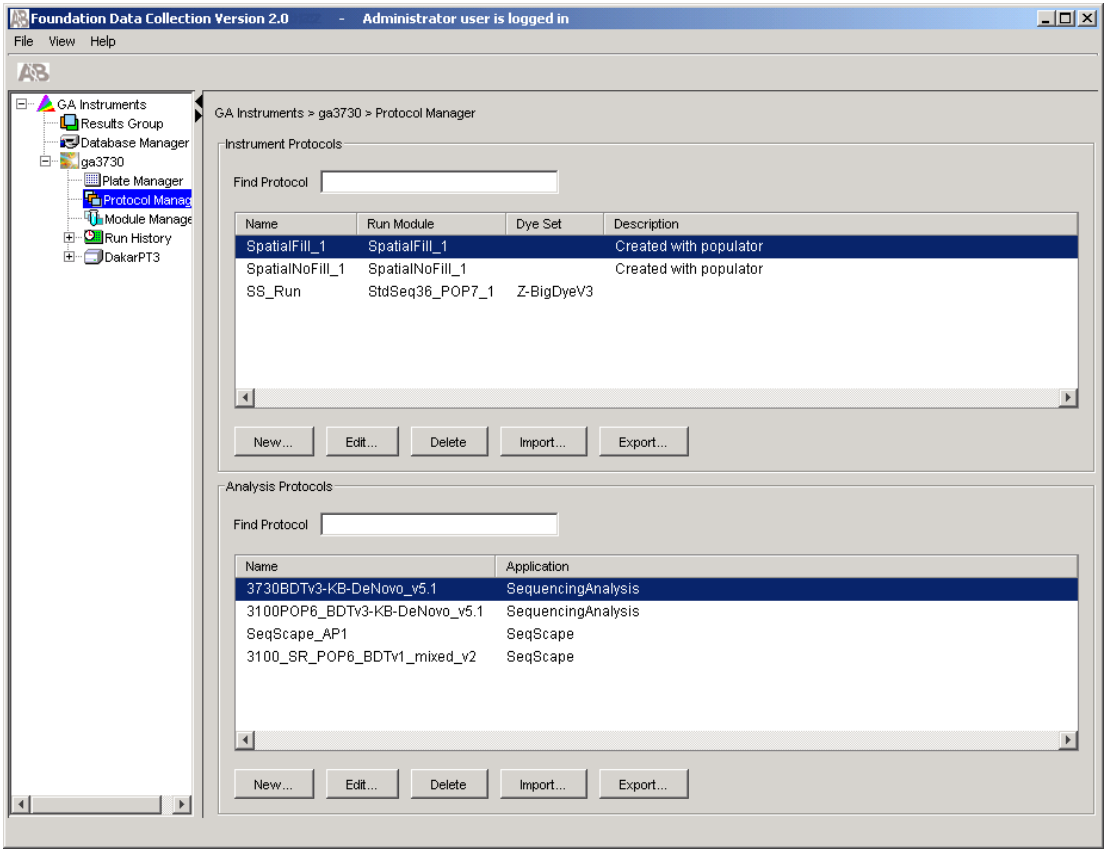
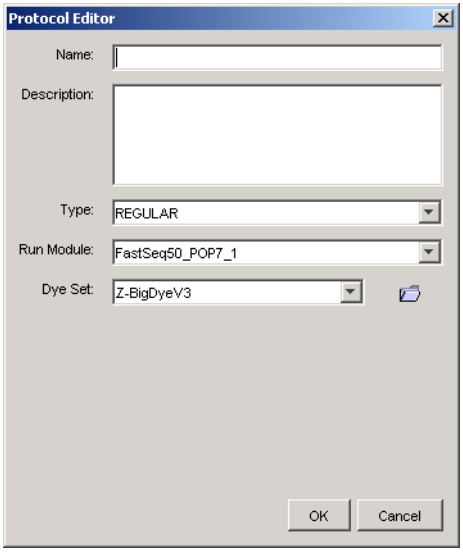


Figure 11-3 Protocol Manager Window

2. In the Instruments Protocols section, click **New** or select an existing instrument protocol file, then click **Edit**.

   The Protocol Editor opens.



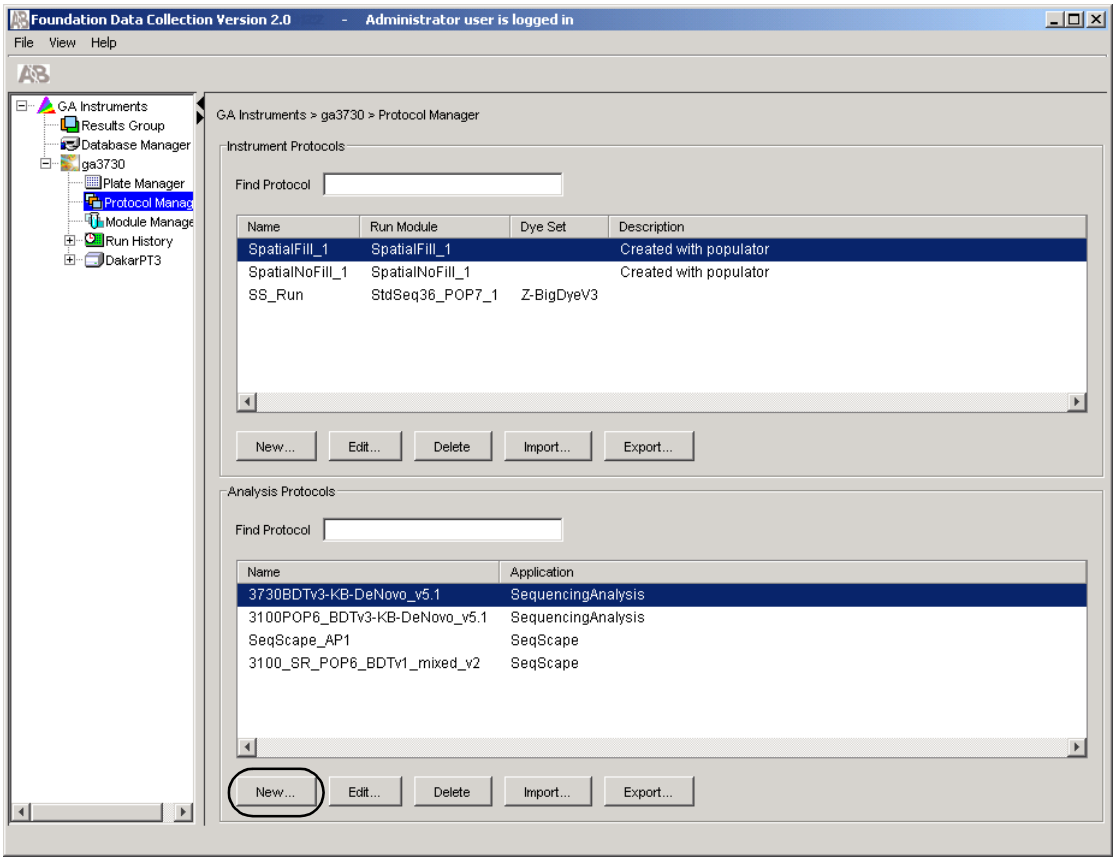
3. Complete the Protocol Editor:
   a. Type a name for the protocol.
   b. Type a description for the protocol (optional).
   c. Select **Regular** in the Type drop-down list.
   d. Select the correct run module for your run. (See the appropriate instrument user guide for information on selecting the correct run module.)
   e. Select the correct Dye Set for your run. (See the appropriate instrument user guide for information on selecting the correct Dye Set.)
   f. Click **OK**.

Creating an Analysis Protocol

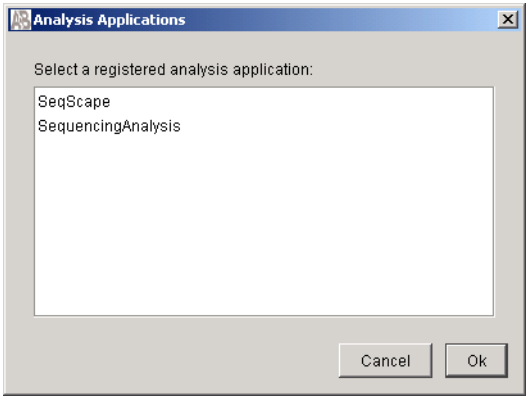
IMPORTANT! If you created an appropriate analysis protocol in SeqScape software, you can use it in data collection software.You can also create an analysis protocol in SeqScape software, if desired.

To create an analysis protocol:

1. In the Analysis Protocols section of the Protocol Manager, click **New**.

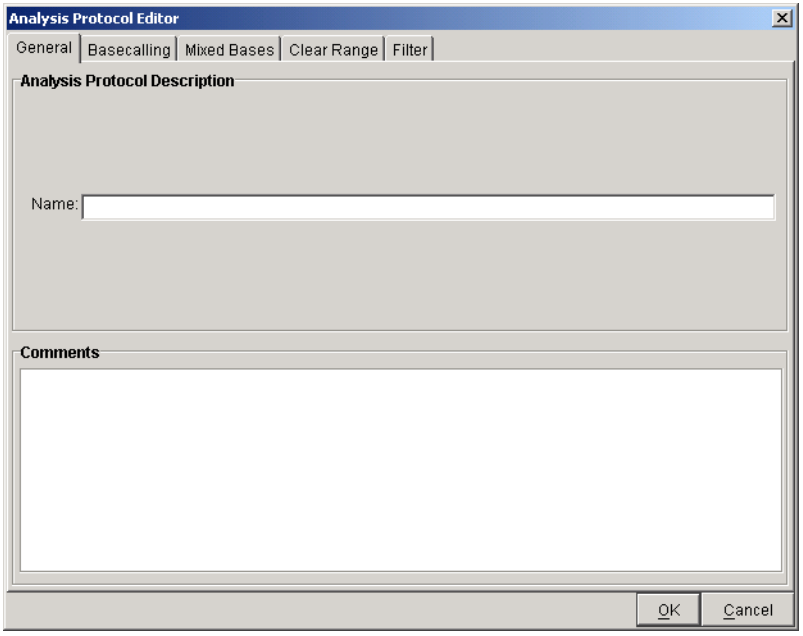


If more than one analysis application is installed on the data collection computer, the Analysis Applications dialog box opens.



2. Select **SeqScape**, then click **OK**.

   The Analysis Protocol Editor dialog box opens.



3. In the General tab, enter a unique name and description for the new protocol.

4. Select the **Basecalling** tab, then do the following:
   a. Select the appropriate basecaller and Dye Set/Primer file. For information on selecting the correct files, see Appendix A, "Basecallers and DyeSet/Primer Files."

      IMPORTANT! Make sure that the basecaller and the DyeSet/Primer file types match.

   b. If desired, select one or more stop points for data analysis. Base your selection on the basecaller being used as indicated below.

Option	Basecaller
At PCR Stop check box	KB or ABI
After __ Ns in __ bases check box	ABI
After __ Ns check box	ABI
After __ Bases check box	KB or ABI

   c. For the KB basecallers only, select True Profile or Flat Profile for processed data and select a quality threshold.

5. Select the **Mixed Bases** tab, then:

   Note: This function is active with the KB basecaller only.

   a. For mixed bases only, select **Use Mixed Base Identification**.
   b. Use the default setting of 25% or change the detection level by entering a new value or dragging the % line up or down.

6. Select the **Clear Range** tab, then, if desired, select one or more stop points for data analysis.

   Note: The clear range is the region of the sequence that remains after excluding the low-quality or error-prone sequence at both the 5´ and 3´ ends.

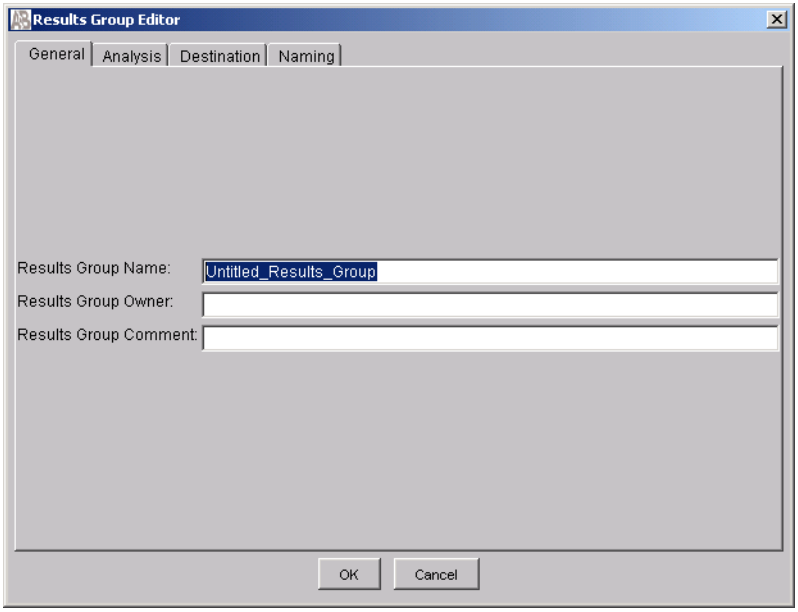
7. Select the **Filter** tab, then, if desired, change one or more of the settings.
8. Select **OK** to save the protocol and close the Analysis Protocol Editor.

Creating a Results Group

To create a results group:

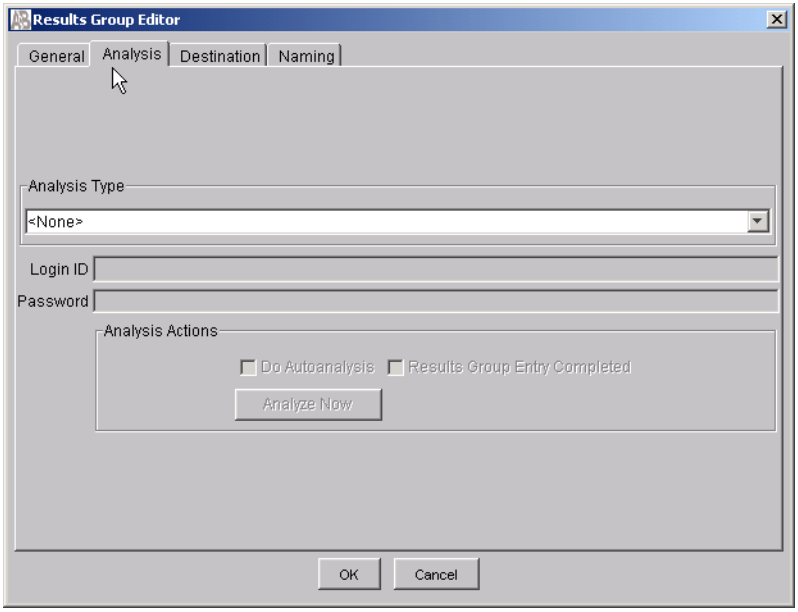
1. Click the **Results Group** icon in the navigation pane.

   The Results Group Editor opens.



2. Click **New**, or select an existing group, then click **Edit**.
3. In the General tab:
   a. Type a Results Group Name. The name can be used in naming and sorting sample files and it must be unique (see “File-Naming Convention” on page 2-10 for a list of accepted characters).
   b. Type a Results Group Owner. The owner name can be used in naming and sorting sample files.
   c. Type a Results Group Comment (optional).

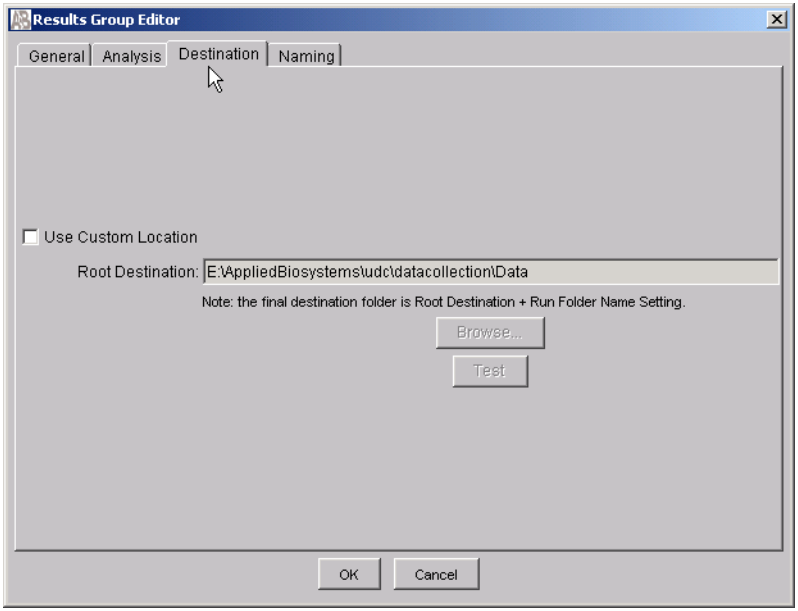
4. Select the **Analysis** tab.



5. In the Analysis tab:
   a. In the Analysis Actions section, select **Do Autoanalysis**.
   b. Select **SeqScape_ *YourInstrumentName*** in the Analysis Type drop-down list.
   c. Type a valid SeqScape Login ID and Password in the text boxes.

IMPORTANT! Failure to use the proper login and password causes your samples not to be analyzed automatically.

6. Select the **Destination** tab.

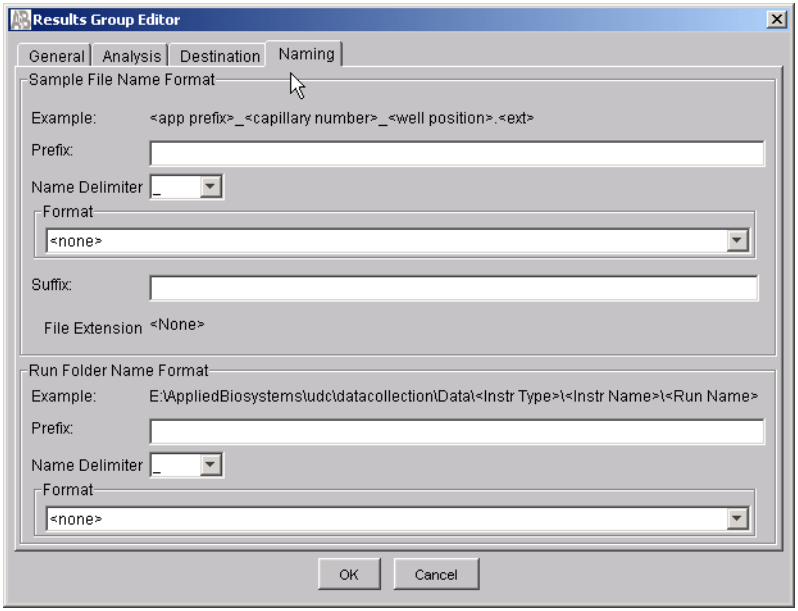


7. In the Destination tab, use the default location or define a new location.
8. To define a new location:
   a. Click **Use Custom Location**, then click **Browse** to navigate to a different save location.
   b. Click **Test** to test the Location path name connection:

      If it passes, a message box displays “Path Name test successful.”

      If it fails, a message box displays “Could not make the connection. Please check that the Path Name is correct.” Click and retry to establish a connection.
   c. Click **OK**.

9. Select the **Naming** tab, then define custom names for sample file and run folder name, if desired.



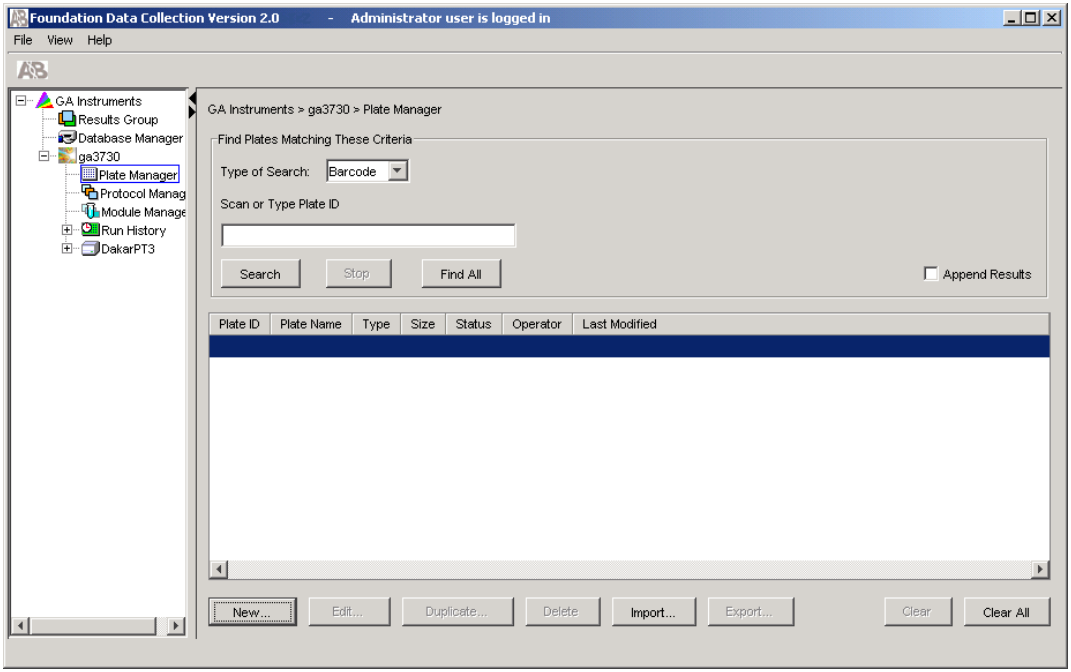
10. Click **OK** to save the Results Group and close the Results Group Editor.

    Note: Specimens in the project can be created automatically based on the delimiter text present in the sample ID name. See “Annotation Tab” on page 6-12.

Creating a Plate Record

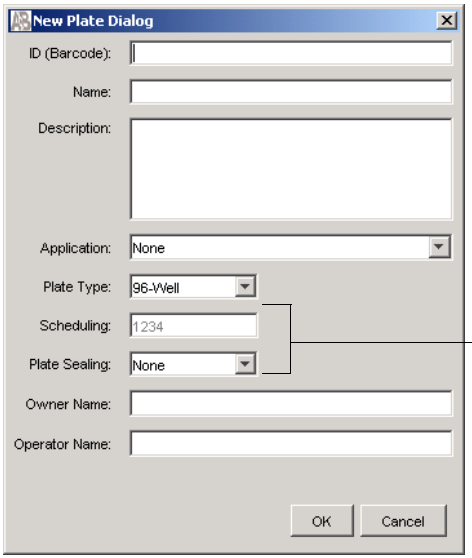
To create a new plate record:

1. Click the **Plate Manager** icon in the navigation pane.



2. Click **New**.

The New Plate Dialog dialog box opens.



3. Complete the information in the New Plate Dialog:
   a. Type an ID for the plate.
   b. Type a name for the plate.
   c. Type a description for the plate (optional).
   d. Select **SeqScape_*YourInstrumentName*** in the Application drop-down list.
   e. Select **96-well** or **384-well** in the Plate Type drop-down list.
   f. For the 3730/3730*xl* instruments, define a scheduling pattern for a 384-well plate.
   g. For the 3730/3730*xl* instruments, select **Heat Sealing** or **Septa** in the Plate Sealing drop-down list.
   h. Type a name for the owner and operator.
   i. Click **OK**.

      The SeqScape Plate Editor opens (see Figure 11-4 on page 11-20).

Completing a Plate Record

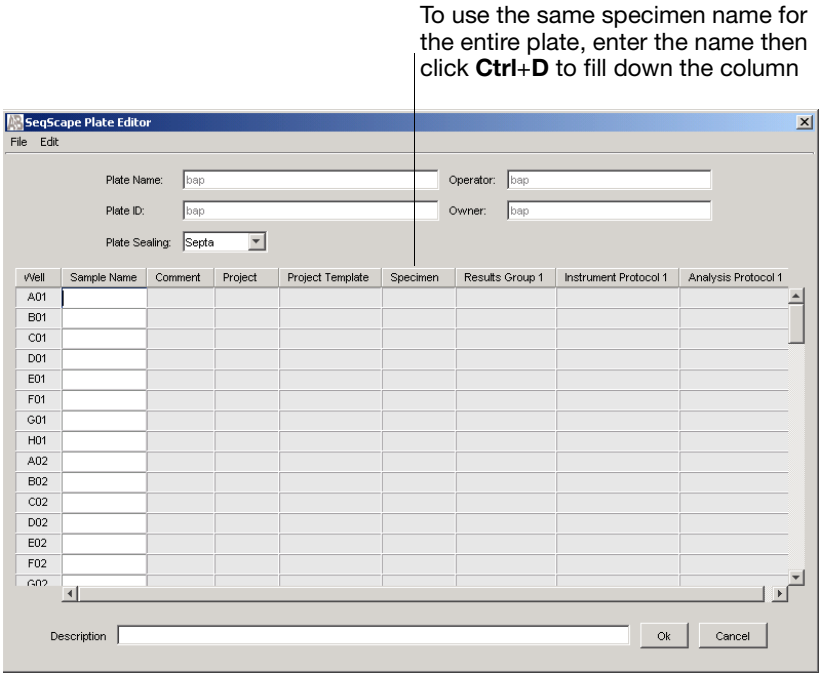


Figure 11-4 Blank Plate Editor

To complete a plate record:

1. Type sample names.

   Note: You cannot use duplicate sample names in the same specimen.

2. Optional: Add comments.

3. Select a project or create a new one.

   If you select a project previously created in SeqScape software, the project template field fills in automatically.

   If you choose to create a new project in the plate editor, you can select only a project template that is already registered in SeqScape software. You *cannot* create a project template in the plate editor.

4. Select a project template.

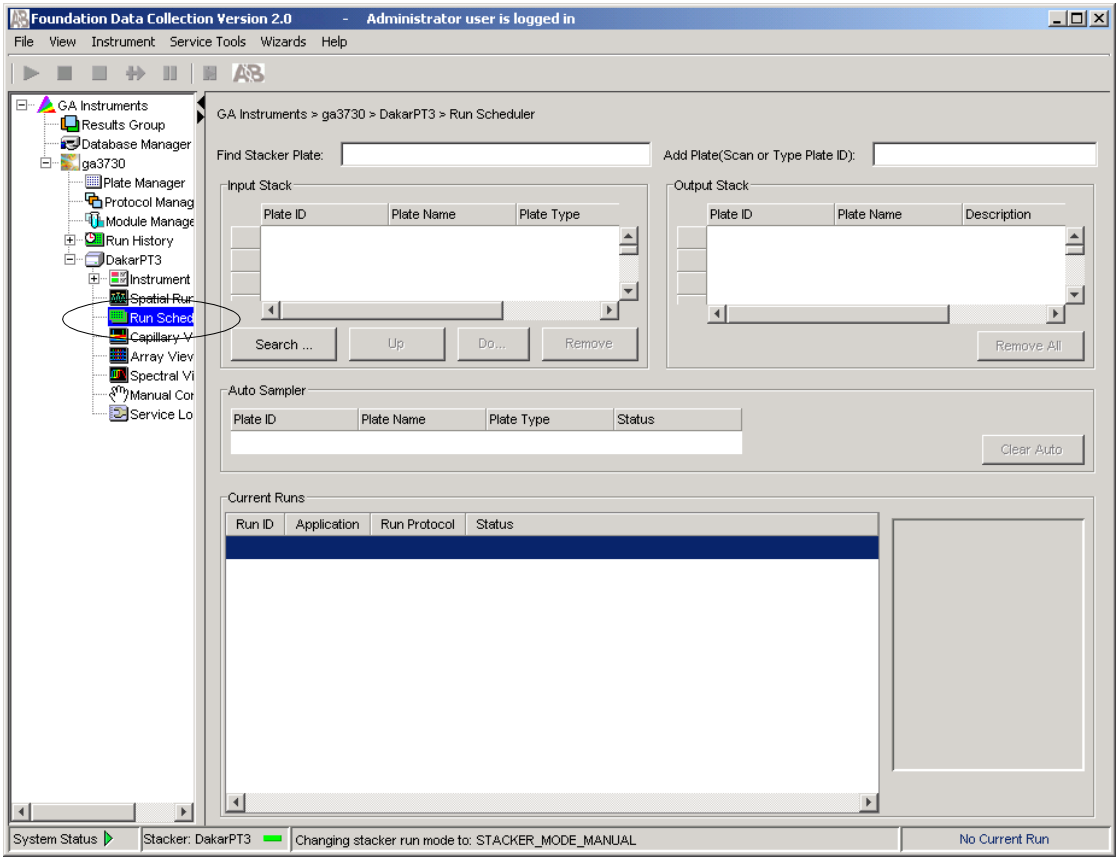
5. Select a specimen or create one.

6. Select a results group or create a new one.
7. Select an instrument protocol or create one.
8. Select an analysis protocol or create one.
9. To add additional runs to the plate record:
    a. Select **Edit > Add Sample Instance**.
    b. Select a results group, instrument protocol, and analysis protocol.
10. Click **OK**.

Scheduling and Starting a Run

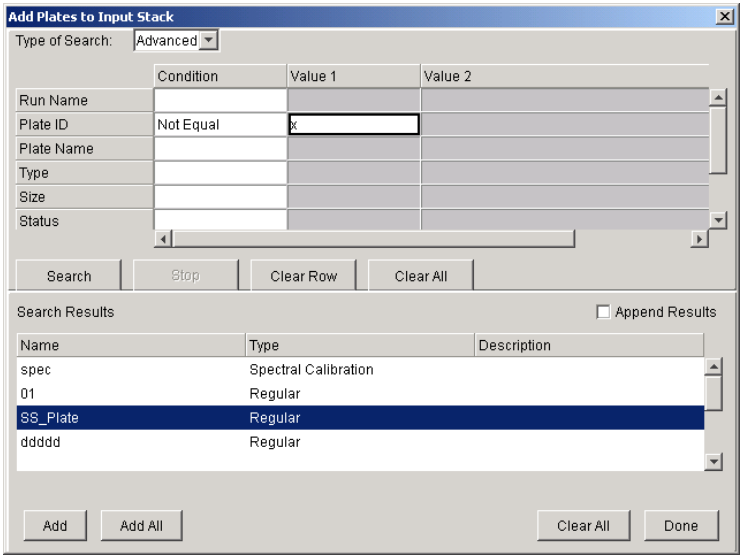
To schedule and then start a run:

1. Click the **Run Scheduler** icon in the navigation pane.



2. Click **Search** in the Input Stack section.

   The Add Plates to Input Stack dialog box opens.



3. Complete the Add Plates to Input Stack dialog box:
   a. Select **Advanced** in the Type of Search drop-down list.
   b. Enter your search criteria for your plate, then click **Search**.
   c. Select the plate in the Search Results section, then click **Add**.
   d. Click **Done**.

      The plate is displayed in the current Runs section of the Run Scheduler window.

4. Click (Run) to start the run.

Understanding the Autoanalysis Manager

Overview

Autoanalysis occurs in the following sequence:

1. When data collection software finishes a run, the Message Service sends the message, “Run Completed.”
2. The Autoanalysis Manager receives the message, and the job is submitted. The job appears in the General tab.
3. The Autoanalysis Manager polls for jobs every 2 minutes and opens the automated processing SeqScape version to analyze the data in the projects.
4. At the end of analysis, the status in the Autoanalysis Manager is updated.

IMPORTANT! Do not run other applications when the Autoanalysis manager and the data collection software are running. Never open a project while it is in use by the Autoanalysis manager.

Files Created

The data collection software stores the sample files in the location specified in the results group. The Autoanalysis Manager copies the files into the DataStore for SeqScape software processing.

To maintain sufficient storage space on your hard drive, delete the sample files created by data collection software that are no longer needed.

Components

The Autoanalysis Manager has two or more tabs:

- General
- SeqScape 2.1
- GeneMapper 3.5 (if loaded)

General Tab

The General tab shows the jobs that have been submitted and their status.

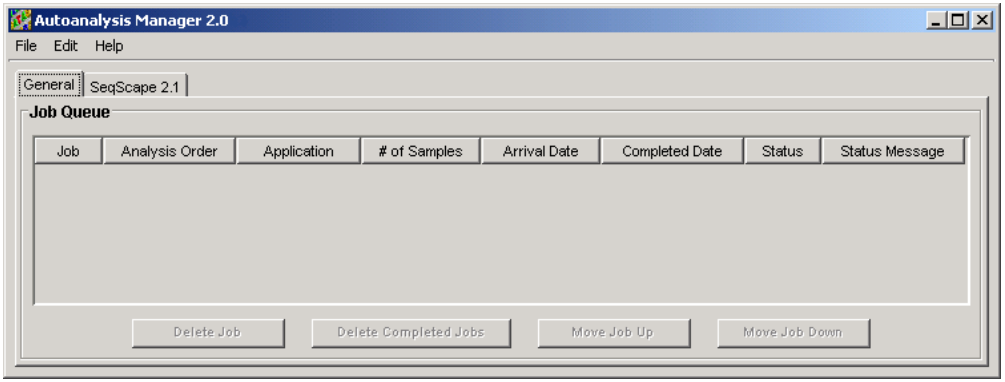


Figure 11-5 Autoanalysis Manager General Tab

Table 11-2 describes the functions of the command buttons in the General tab.

Table 11-2 General Tab Command Buttons

Button Name	Function
Delete Job	Deletes a pending job
Delete Completed Jobs	Deletes a completed job
Move Job Up	Moves a pending job higher in the queue
Move Job Down	Moves a pending job lower in the queue

SeqScape 2.1 Tab

The SeqScape 2.1 tab shows the jobs, project, and status information.

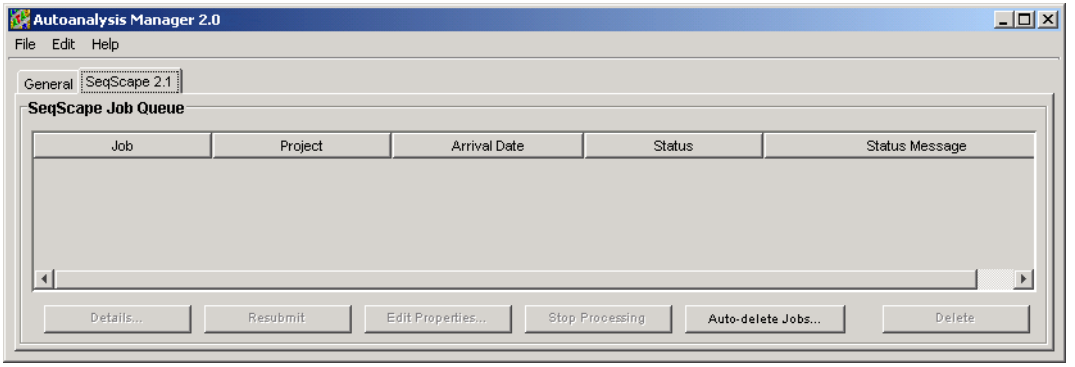


Figure 11-6 Autoanalysis Manager SeqScape 2.1 tab

Table 11-3 describes the functions of the command buttons in the SeqScape 2.1 tab:

Table 11-3 SeqScape 2.1 Tab Command Buttons

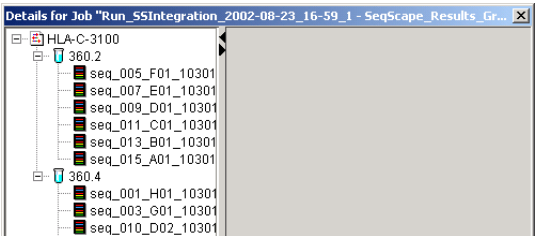
Button Name	Function
Details	Displays the project in the navigation pane 
Resubmit	If a job fails, resubmit for analysis
Edit Properties	Edits the name and password (active only if analysis failed)
Stop Processing	Interrupts the processing of the current job this. The job will be marked as “failed” and the auto analysis manager will proceed to the next job.

Table 11-3 SeqScape 2.1 Tab Command Buttons

Button Name	Function
Auto-delete Jobs	Automatically deletes the completed jobs from the auto analysis manager. This option allows you to select when to delete, in number of days, the completed jobs. This option is a convenient way to save disk space.
Delete	Deletes a job from the Autoanalysis Manager

File Sharing Between Data Collection and SeqScape Software

In Table 11-4, the term “files” refers to projects, project templates, specimens, and analysis protocols.

Table 11-4 File Sharing Table

Conditions	Result	Corrective Action
SeqScape software installed while data collection software was open (proper installation)		
• Files created in SeqScape software • Data collection software open	Files are registered in both applications and are available for use in the data collection software.	—
• Files created in data collection software • SeqScape software open	Files are registered in both applications and are available for use in SeqScape software.	—
SeqScape software installed while data collection software was closed (improper installation)		
• Files created in SeqScape software or in data collection software • Other software open or closed	SeqScape software was never registered in the Data Service—no communication between the software.	1. Uninstall SeqScape software. 2. Open the data collection software. 3. Reinstall SeqScape software. 4. Register the software and define user IDs.

Basecallers and DyeSet/Primer Files

A

This appendix contains:

Definitions and Naming A-2
ABI PRISM 310 Genetic Analyzer Files A-5
ABI PRISM 377 DNA Sequencer Files A-8
ABI PRISM 3100 Genetic Analyzer Files A-10
ABI PRISM 3100-*Avant* Genetic Analyzer Files A-13
ABI PRISM 3700 DNA Analyzer Files A-15
Applied Biosystems 3730/3730xl DNA Analyzers Files A-17

Definitions and Naming

Basecaller Algorithm

A basecaller is an algorithm that determines the bases within a sequence during analysis. There are two types of basecallers:

- KB basecaller – An algorithm that calculates mixed or pure bases, and sample quality values.
- ABI basecaller – An algorithm used in earlier versions of ABI PRISM® Sequencing Analysis and ABI PRISM® SeqScape® Software.

DyeSet/Primer File

The DyeSet/Primer file compensates for the mobility differences between the dyes and primers and corrects the color code changes due to the type of chemistry used to label the DNA. DyeSet/Primer files are sometimes referred to as mobility files.

DyeSet/Primer File-Naming Conventions

DyeSet/Primer files use the following name convention:

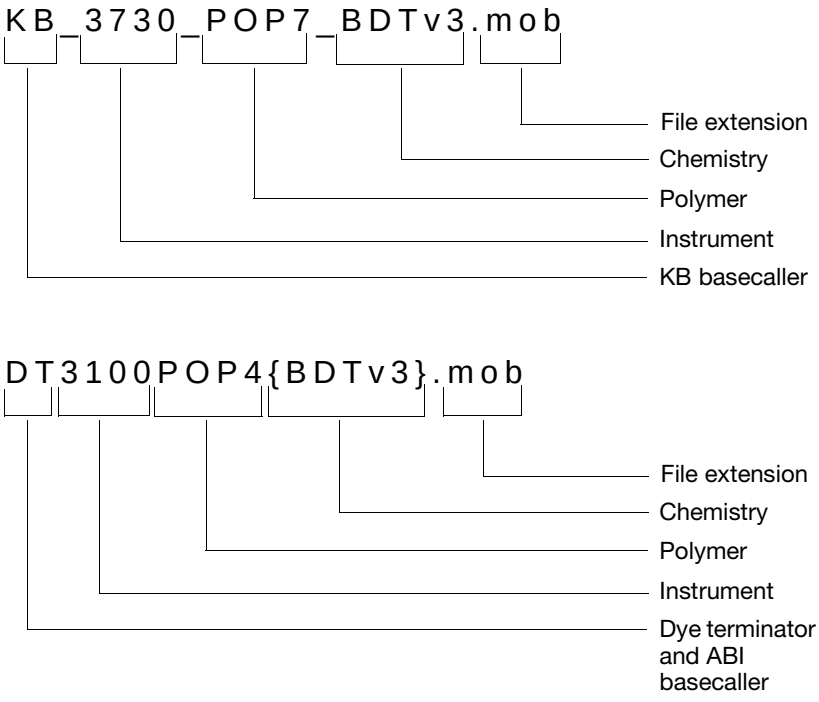


Figure A-1 Examples of DyeSet/Primer File Naming Convention

The DyeSet/Primer file names use a combination of characters to indicate the basecaller, instrument, chemistry, and polymer type as described in Table A-1.

Table A-1 DyeSet/Primer File Names

Abbreviation	For Runs Using …
Basecaller	
KB	KB basecaller
DP	Dye primer chemistry and the ABI basecaller
DT	Dye terminator chemistry, and the ABI basecaller
Type of Polymer or Gel	
5%LR	% Long Ranger in the gel (377 instrument only)
POP4	ABI PRISM® POP-4™ polymer
POP5	ABI PRISM® POP-5™ polymer
POP6	ABI PRISM® POP-6™ polymer
POP7	ABI PRISM® POP-7™ polymer
Chemistry	
BDTv3	ABI PRISM® BigDye® v3.0 and 3.1 Terminator chemistry
{BDv3}	
{BDv1}	ABI PRISM® BigDye® v1.0 and 1.1 Terminator chemistry
{BD}	
{-21M13}	Dye primer chemistry – the -21M13 primer is labeled
{M13Rev}	Dye primer chemistry – the M13Rev primer is labeled

Basecaller and DyeSet/Primer Compatibility

The DyeSet/Primer file must match the chemistry and basecaller type that you are using

Note: DyeSet/Primer files are filtered based on the selected basecaller.

IMPORTANT! However, if you select the DyeSet/Primer file then select a basecaller file, no filtering of the basecaller list occurs. If you select a KB DyeSet/Primer file and an ABI basecaller for analysis, or a DT DyeSet/Primer file and an KB basecaller for analysis, the following error dialog box opens (see Figures A-2 and A-3).

BaseCaller - DyeSet/Primer Incompatibility
Please choose a compatible DyeSet/Primer file for the BaseCaller which you have chosen.
OK

Figure A-2 Error Message in the Sample Manager

Figure A-3 Error Message in the Analysis Protocol

ABI PRISM 310 Genetic Analyzer Files

Note: 47 cm capillary array length = 36 cm read length
61 cm capillary array length = 50 cm read length

Table A-2 310 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM® BigDye® Terminator v1.0 and v1.1	47	KB.bcp	KB_310_POP4_BDTv1_36Rapid.mob KB_310_POP4_BDTv1_36Std.mob
	47	KB.bcp	KB_310_POP6_BDTv1_36Rapid.mob
	61		KB_310_POP6_BDTv1_50Std.mob
ABI PRISM® BigDye® Terminator v3.0 and v3.1	47	KB.bcp	KB_310_POP4_BDTv3_36Rapid.mob KB_310_POP4_BDTv3_36Std.mob
	47	KB.bcp	KB_310_POP6_BDTv3_36Rapid.mob
	61		KB_310_POP6_BDTv3_50Std.mob
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	47	Basecaller-310POP4.bcp	DT310POP4{BD}v2.mob
	47	Basecaller-310POP6.bcp	DT310POP6{BD}.mob
	61		DT310POP6{BD-LR}v3.mob

Table A-2 310 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry *(continued)*

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI PRISM® dRhodamine Terminator	47	Basecaller-310POP4.bcp	DT310POP4{dRhod}v1.mob
	47	Basecaller-310POP6.bcp	DT310POP6{dRhod}v2.mob
	61		
ABI PRISM BigDye Terminator v3.0 and v3.1	47	Basecaller-310POP4.bcp	DT310POP4{BDv3}v2.mob
	47	Basecaller-310POP6.bcp	DT310POP6{BDv3}v2.mob
	61		

Table A-3 310 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecaller			
ABI PRISM® BigDye® Primer v1.0 and v1.1	47	Basecaller-310POP4.bcp	DP310POP4{BD-21M13}v1.mob DP310POP4{M13Rev}v1.mob
	47	Basecaller-310POP6.bcp	DP310POP6{BD-21M13}v1.mob DP310POP6{M13Rev}v1.mob
	61		

Table A-3 310 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry *(continued)*

ABI PRISM BigDye® Primer v3.0 and v3.1	47	Basecaller-310POP4.bcp	DP310POP4{BDv3-21M13}v1.mob DP310POP4{BDv3-M13Rev}v1.mob
	47	Basecaller-310POP6.bcp	DP310POP6{BDv3-21M13}v1.mob DP310POP6{BDv3-M13Rev}v1.mob
	61		

ABI PRISM 377 DNA Sequencer Files

Table A-4 377 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	WTR (cm)/Scan Rate (scans/hr)	Basecaller	DyeSet/Primer
ABI Basecalling			
• ABI PRISM BigDye Terminator v1.0 and v1.1 • ABI PRISM dGTP BigDye Terminator	36/2400	Basecaller-377.bcp	DT377{BD}.mob
	36 & 48/1200	Basecaller-377LR.bcp	
ABI PRISM dRhodamine Terminator	36/2400	Basecaller-377.bcp	DT377{dRhod}.mob
	36 & 48/1200	Basecaller-377LR.bcp	
• ABI PRISM BigDye Terminator v3.0 and 3.1 • ABI PRISM dGTP BigDye v3.0 Terminator	36/2400	Basecaller-377.bcp	DT377{BDv3}v2.mob
	36 & 48/1200	Basecaller-377LR.bcp	DT377LR{BDv3}v1.mob

Table A-5 377 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	WTR (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Primer v1.0 and v1.1	36/2400	Basecaller-377.bcp	DP377-5%LR{BD-21M13}.mob DP377-5%LR{BD-M13Rev}.mob,
	36 & 48/1200	Basecaller-377LR.bcp	
ABI PRISM BigDye Primer v3.0 and 3.1	36/2400	Basecaller-377.bcp	DP377{BDv3-21M13}v1.mob DP377{BDv3-M13Rev}v1.mob
	36 & 48/1200	Basecaller-377LR.bcp	

ABI PRISM 3100 Genetic Analyzer Files

Table A-6 3100 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv1.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv1.mob
	50: std read		
ABI PRISM BigDye Terminator v3.0 and v3.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv3_.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv3.mob
	50: std read		

Table A-6 3100 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry *(continued)*

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
• ABI PRISM BigDye Terminator v1.0 and v1.1 • ABI PRISM dGTP BigDye Terminator	36: ultra rapid	Basecaller-3100POP4UR.bcp	DT3100POP4LR{BD}v1.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100POP6RRv2.bcp	DT3100POP6{BD}v2.mob
	50: std read	Basecaller-3100POP6SR.bcp	
• ABI PRISM BigDye Terminator v3.0 and 3.1 • ABI PRISM dGTP BigDye v3.0Terminator	36: ultra rapid	Basecaller-3100POP4UR.bcp	DT3100POP4{BDv3}v1.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100POP6RRv2.bcp	DT3100POP6{BDv3}v1.mob
	50: std read	Basecaller-3100POP6SR.bcp	
ABI PRISM dRhodamine Terminator	36: ultra rapid	Basecaller-3100POP4UR.bcp	DT3100POP4{dRhod}v2.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100POP6RRv2.bcp	DT3100POP6{dRhod}v2.mob
	50: std read	Basecaller-3100POP6SR.bcp	

Table A-7 3100 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Primer v1.0 and v1.1	36: rapid read	Basecaller-3100POP6RRv2.bcp	DP3100POP6{BD-21M13}v1.mob DP3100POP6{BD-M13Rev}v1.mob
	50: std read	Basecaller-3100POP6SR.bcp	
ABI PRISM BigDye Primer v3.0 and 3.1	36: rapid read	Basecaller-3100POP6RRv2.bcp	DP3100POP6{BDv3-21M13}v1.mob DP3100POP6{BDv3-M13Rev}v1.mob
	50: std read	Basecaller-3100POP6SR.bcp	
ABI PRISM BigDye v3 Primer (All primers)	36: ultra rapid	Basecaller-3100POP4UR.bcp	DP3100POP4{BDv3}v1.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	

ABI PRISM 3100-*Avant* Genetic Analyzer Files

Table A-8 3100-*Avant* Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv1.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv1.mob
	50: std read		
ABI PRISM BigDye Terminator v3.0 and v3.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv3_.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv3.mob
	50: std read		

Table A-8 3100-*Avant* Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry *(continued)*

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: ultra rapid	Basecaller-3100APOP4UR.bcp	DT3100POP4LR{BD}v1.mob
	80: long read	Basecaller-3100APOP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100APOP6RRv2.bcp	DT3100POP6{BD}v2.mob
	50: std run	Basecaller-3100APOP6SR.bcp	
ABI PRISM BigDye Terminator v3.0 and 3.1	36: ultra rapid	Basecaller-3100APOP4UR.bcp	DT3100POP4{BDv3}v1.mob
	80: long read	Basecaller-3100APOP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100APOP6RRv2.bcp	DT3100POP6{BDv3}v1.mob
ABI PRISM dRhodamine Terminator	36: ultra rapid	Basecaller-3100APOP4UR.bcp	DT3100POP4{dRhod}v2.mob
	80: long read	Basecaller-3100APOP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100APOP6RRv2.bcp	DT3100POP6{dRhod}v2.mob
	50: std run	Basecaller-3100APOP6SR.bcp	

ABI PRISM 3700 DNA Analyzer Files

Table A-9 3700 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and 1.1	50	Basecaller-3700POP6.bcp	DT3700POP6{BD}v5.mob
		Basecaller-3700POP5LR.bcp	DT3700POP5{BD}v3.mob
ABI PRISM BigDye Terminator v3.0 and 3.1	50	Basecaller-3700POP6.bcp	DT3700POP6{BDv3}v1.mob
		Basecaller-3700POP5LR.bcp	DT3700POP5{BDv3}v1.mob
ABI PRISM dRhodamine Terminator	50	Basecaller-3700POP6.bcp	DT3700POP6{dRhod}v3.mob
		Basecaller-3700POP5LR.bcp	DT3700POP5{dRhod}v1.mob

Table A-10 3700 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Primer v1.0 and v1.1	50	Basecaller-3700POP6.bcp	DP3700POP6{BD-21M13}v3.mob DP3700POP6{BD-M13Rev}v2.mob
		Basecaller-3700POP5LR.bcp	DP3700POP5{BD-21M13}v1.mob DP3700POP5{BD-M13Rev}v1.mob
ABI PRISM BigDye Primer v3.0 and v3.1	50	Basecaller-3700POP6.bcp	DP3700POP6{BDv3-21M13}v1.mob DP3700POP6{BDv3-M13Rev}v1.mob
		Basecaller-3700POP5LR.bcp	DP3700POP5{BDv3-21M13}v1.mob DP3700POP5{BDv3-M13Rev}v1.mob

Applied Biosystems 3730/3730*xl* DNA Analyzers Files

Table A-11 3730/3730*xl* Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM BigDye v3.0 Terminator	all lengths	KB.bcp	KB_3730_POP7_BDTv3.mob
ABI PRISM BigDye Terminator	all lengths	KB.bcp	KB_3730_POP7_BDTv1.mob
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: rapid read	Basecaller-3730POP7RR.bcp	DT3730POP7{BD}.mob
	36: std read	Basecaller-3730POP7SR.bcp	
	50: long read	Basecaller-3730POP7LR.bcp	
ABI PRISM BigDye Terminator v3.0 and v3.1	36: rapid read	Basecaller-3730POP7RR.bcp	DT3730POP7{BDv3}.mob
	36: std read	Basecaller-3730POP7SR.bcp	
	50: long read	Basecaller-3730POP7LR.bcp	

Frequently Asked Questions B

This appendix contains:

Upgrading FAQs B-2
Training and Documentation FAQs. B-3
SeqScape Software Basics FAQs B-4
General SeqScape Software FAQs B-5
SeqScape Manager FAQs B-9
Library FAQs. B-15
Mutation, Variant, HIM, and HFM Detection FAQs. B-16
Data Analysis FAQs B-17
Analysis Reports FAQs B-19
Quality Values FAQs B-20
Printing and Exporting Results FAQs B-21
Audit Trail, Security, and Access Control FAQs. B-22

Upgrading FAQs

There are four versions of ABI PRISM® SeqScape® Software v2.1 available. See “SeqScape Software Versions” on page 2-5.

Table B-1 Upgrading Questions and Answers

Question	Answer
How does SeqScape software v2.1 differ from v2.0?	Refer to Chapter 1, “Introduction to ABI PRISM SeqScape Software.”
What happens to my data when I upgrade my SeqScape software? What happens to my data if I do an install of the SeqScape software v2.1?	If you install the upgrade version of SeqScape software version 2.1, data from SeqScape software v1.0, v1.1, or 2.0 will be stored in the Data Store backup folder that is named according to the version of SeqScape software previously installed (i.e., Data v1.0 or Data v1.1). In addition, all previous contents of the Data Store will be copied to the new Data Store. If you install the full version of SeqScape software v2.1, you must first export the data objects from the SeqScape Manager in the previous version of SeqScape software, then import them into SeqScape software v2.1.
What happens to the data that I created with the demo version when I upgrade to the full version of SeqScape software?	You need to export the data generated with the 30-day demo before it expires, install SeqScape software v2.1, then import this data into the full version of the SeqScape software v2.1. To export, select the desired object from the SeqScape Manager window and click **Export**. To import the object into the full copy of the SeqScape software v2.1, use the SeqScape Manager.

Training and Documentation FAQs

Table B-2 Training and Documentation Questions and Answers

Question	Answer
How do I train myself on SeqScape software?	The best way to train yourself on SeqScape software is to use the training movie and the software tutorial included in the software package. A printed tutorial is included with all SeqScape software versions except for the demo version. The tutorial is also available online, see below.
Where can I get find resources/documentation on SeqScape software?	To find documentation on SeqScape software, go to www.appliedbiosystems.com > **Support** > **Product and Service Literature**. In the Document Title/Keyword dialog box at the bottom of this page enter “SeqScape” and click **Search**.

SeqScape Software Basics FAQs

Table B-3 SeqScape Software Basics Questions and Answers

Question	Answer
What is SeqScape software?	SeqScape software is a resequencing software tool designed to identify nucleotide variants, amino acid variants, and sequences in a library that match each consensus sequence.
SeqScape software can be used for which sequencing application?	SeqScape software can be used for SNP discovery and validation, mutation analysis and heterozygote identification, sequence confirmation for mutagenesis or clone-construct confirmation studies, and the identification of genotype, allele and haplotype from a library of known sequences.
What is the overall workflow for analyzing and reviewing data?	All analysis in SeqScape software occurs in a project. Analysis and review of the data occur in six steps as follows: 1. Create an analysis defaults and display settings (or use one previously created). 2. Create an RDG (or use one previously created). 3. Create a project template (or use one previously created). 4. Create and analyze a project by adding sample files to a project template. 5. Review the results (view/edit the data and view the reports). 6. Export/print the results and reports.
What happens in the SeqScape software when I click (Analyze)?	Each time you click the Analyze button, the following pipeline is applied to specimens that have not been analyzed: 1. Basecall and determine Quality Values 2. Identify mixed bases. (This step may be bypassed if data was previously basecalled.) 3. Trim low quality sequence ends 4. Filter (omit poor quality sequences) 5. Assemble sequences to the reference 6. Generate a consensus sequence for each specimen 7. Align consensus sequence to the reference sequence 8. Compare each consensus to the reference 9. Search the allele Library for matches to each consensus sequence. (This step may be bypassed if you do not have a need to identify allele matches.) 10. Generate Reports

General SeqScape Software FAQs

Table B-4 General Questions and Answers

Question	Answer
Instruments — What Applied Biosystems instruments can I use to generate data for SeqScape software?	SeqScape software analyzes sequence files generated from ABI PRISM 3700, 3100, 3100-*Avant*, 377, 310, and Applied Biosystems 3730 and 3730*xl* genetic analyzers. The software also accepts text sequences in FASTA format.
Instruments — Can SeqScape software be used to analyze data that was generated on instruments other than ABI PRISM or Applied Biosystems instruments?	No. Sequencing data generated on platforms other than Applied Biosystems platforms are not compatible with SeqScape software.
Sample files — Can SeqScape software be used for autoanalysis of sample files?	SeqScape software v2.1 can be used to autoanalyze sample files from Applied Biosystems 3730/3730xl DNA Analyzers and ABI PRISM 3100/3100-*Avant* Genetic Analyzers. For more detailed information on how to configure the software for the autoanalysis of these files, see Chapter 11, “Automating Analysis.”
Sample files — What instrument sample files can I use with the KB basecaller?	You can use the KB basecaller in SeqScape software to analyze sequencing sample files generated from the ABI PRISM 3100 and 3100-Avant genetic analyzers and Applied Biosystems 3730/3730xl DNA Analyzers.
How can I share my work with someone at a different site? What should I send them?	All sample files, analysis parameters, reference sequence, and analysis results are saved in every SeqScape project file. You can share these files with anyone who has the software by exporting the project or data objects and importing them into their software on their computer. There is no link between different SeqScape software installed on different computers. You can also share project templates, which contain the reference sequence and analysis parameters. A colleague can then analyze sample files of their choice using the project templates to create a new project. The analysis is identical to your own analysis with the same project template.

Table B-4 General Questions and Answers *(continued)*

Question	Answer
Files — What are the file types/formats accepted by SeqScape software?	**Reference Sequences** — SeqScape software accepts the following file formats for reference sequences: • Genbank File Format (with a .fcgi, .cgi, or .gb extension) • .txt (text) file format • .ab1 file format • .fsta (FASTA) file format • .seq file format • Aligned sequences in .fsta (FASTA) format* *The imported sequence will show a summary of all the sequences in the file by substituting the IUPAC codes for bases where there is a discrepancy in the sequences. **Data Sequences**— SeqScape software accepts the following file formats for data sequences used for analysis: • .ab1 files (previously basecalled or not basecalled) • txt (TEXT) files • .fsta (FASTA) files • .seq files **Nucleotide Variants** — SeqScape software accepts the following file formats for nucleotide variants: • .fsta file containing a set of aligned sequences in FASTA format. • tab-delimited text (.txt) file that lists one variant per line and eight column headings: Type, ROI, NT position, Reference, Variant, Style, Description, and Used by all ROIs. **Amino Acid Variants** — SeqScape software accepts the following file formats for amino acid variants: • tab-delimited text (.txt) file that lists one variant per line and the following seven column headings: Type, Layer, AA position, Reference, Variant, Style, and Description
Files — Can .scf files be analyzed in SeqScape software?	No. SeqScape software analyzes only sequencing data in .ab1 files or text sequences.
Files — How can I use sample files generated on the Macintosh® computer with SeqScape software?	To use data files generated on a Macintosh computer with SeqScape software, you must convert the files using the SCU (Sample Conversion Utility). This utility is available as a Macintosh application file on the SeqScape Analysis software CD. The SCU must be loaded onto and launched from a Macintosh computer. For more information, see the Read Me file associated with the SCU.

Table B-4 General Questions and Answers *(continued)*

Question	Answer
Chemistry — What Applied Biosystems chemistries are supported?	• ABI PRISM® BigDye® Terminator v3.1, v3, v2, v1.1, and v1 chemistries • ABI PRISM® BigDye® Primers and dRhodamine dyes
Computer — What are the computer requirements for SeqScape software?	• CPU – 733 MHz or faster, single processor • Memory – 512 MB RAM • OS – Windows® 2000 with Service Pack 3, or Windows® XP with Service Pack 1 • 1 GB hard drive • Pentium® III or IV chip, *not* Xenon
Computer — What kind of performance can I expect from my SeqScape software?	That depends on your computer specifications. For example, if you have a computer with an 850 MHz processor, 256 MB RAM, Pentium® III chip, running on Windows 2000 OS and are analyzing 100 samples files, 10 specimens, and a 1Kb reference, the analysis time is 2 min. **Note:** If SeqScape software is running slowly, you can improve performance by archiving older projects. To archive projects, select **Tools** > **SeqScape Manager** > **Project** > **Export**.
Software — Does SeqScape software support BioLIMS/Sequence Collector software?	No. SeqScape software no longer supports BioLIMS/Sequence Collector software.
Software — How does SeqScape software compare to MicroSeq® and ViroSeq™ software?	SeqScape software – compares samples to a reference sequence MicroSeq software – identifies bacteria ViroSeq software – identifies genotype HIV-1 resistance mutations
Software — Do I need Sequencing Analysis software if I have SeqScape software?	Sequencing Analysis software is a multi-purpose software used to analyze, edit, view, display, and print sequencing sample files. SeqScape software is designed specifically for resequencing. Sequencing Analysis software should be used in every laboratory for general troubleshooting and viewing of data.
Data objects— Can I transfer data objects like the RDG, Display Settings, Analysis Protocols, etc. from one computer to another?	Yes, you can transfer data objects from one computer to another. To export the data object, to the SeqScape Manager, select the object and click **Export**. Send the exported file to the second computer and launch SeqScape software. Open SeqScape Manager and click **Import** to import the file.

Table B-4 General Questions and Answers *(continued)*

Question	Answer
Can I BLAST against a database?	To search a database using a sequence generated with SeqScape software, in the Project view, export the NT alignment as an aligned FASTA file by selecting **File** > **Export**. Open this file in a text viewer, then cut and paste the sequence you would like to search for in your BLAST query. Refer to Chapter 9, "Exporting and Printing Data and Reports," for detailed information on exporting.
Alignment – What alignment algorithms are used in SeqScape software?	The sample assembly and specimen alignments are generated using a Smith-Waterman local sequence alignment algorithm using parameters appropriate for DNA sequencing.
Alignment – Can SeqScape software perform just the alignment for samples?	Yes. In order to assemble and analyze sequences without basecalling, open the Analysis Defaults for the project and select the **Specimen** tab. Deselect the **Basecall Samples** check box.

SeqScape Manager FAQs

Table B-5 SeqScape Manager Questions and Answers

Question	Answer
What is the SeqScape Manager?	SeqScape Manager allows you to import, export, create, and delete projects, project templates, reference data groups, analysis defaults, libraries, analysis protocols, and display settings. Access SeqScape Manager by selecting it from the Tools menu.
What is an object	An object is a named collection of data elements to perform certain functions, for example, analysis protocol.
How do I create a new user?	You must be logged in as Admin user. 1. Select **Tools** > **Options**. 2. Select the **Users** tab, then click **New**. 3. Enter the new user name (be sure to omit any spaces in the user name), then click **OK**. 4. To log in with the new name, exit the software, then relaunch it. 5. Log in with the new user name.
What is a project in SeqScape software?	A project is created using a project template. Projects contain sample data files grouped into specimens.
What is a project template?	A project template is the mold from which projects are created. Templates contain: analysis defaults, display settings, and a reference data group.
What is a specimen?	A specimen contains all the sample data from a single biological source. SeqScape software will assemble all sample data within a specimen and generate a consensus sequence. For example, a specimen contains forward strand PCR products for exons 3, 4 and 5 of a gene and several reverse strand PCR products for the same exons. The software will generate a single consensus sequence representing exons 3,4 and 5 and compare it to the reference sequence. Do not mix products from different biological sources into a single specimen.
What is a layer?	A layer is a set of ROIs that are grouped together for the purpose of display, report or amino acid translation. The ROIs within a layer cannot overlap. Example: Your project may contain introns 1,2,3, 4,5,6 and exons 1,2,3,4,5. You may choose to create a layer that contains only exons 1,4,5 or a layer containing intron1, exon1,intron2, exon2, intron3, exon3, and intron4. A layer may represent a transcript.

Table B-5 SeqScape Manager Questions and Answers *(continued)*

Question	Answer
How do I generate a new layer?	You can generate a new layer by clicking on **New Layer** in the ROI tab of the RDG. Then choose to add the desired ROIs by checking the **On Layer** check box in the ROI table. Also, select whether you would like the ROI translation turned in the layer.
Can I put samples from different individuals in the same specimen?	No. Each individual sample should be in a different specimen. Refer to Chapter 6.
Can I mix samples from different biological sources?	It is not possible to analyze data from different biological sources in the same specimen.
What is a reference data group (RDG)?	The RDG is an essential part of the project template that contains all of the analysis-specific information, including the reference sequence, translation codon table, known variants, RDG name, reference segments, regions of interest (ROI), layers, and the name of the associated allele libraries.
What are the new features of the extended RDG?	When using a Genbank file to create the reference sequence, the feature table of the Genbank file is pulled into the RDG, and each feature is listed out in the ROI (region of interest) table. These ROIs can be used to create new layers for sequence comparison. For example, if a Genbank file for a gene containing two exons and one intron is imported into the RDG, the user can choose to create a layer that includes only the two exons. When analysis occurs, the specimens will be compared to this layer containing the two exons as well as the reference backbone layer that includes the two exons and the intron. In addition, the user can choose to turn translation on or off for specific ROIs. A library containing aligned sequences can also be attached to a specific layer for comparison during analysis.
What is a reference sequence?	A reference sequence is the backbone sequence against which the software compares the consensus segments. A reference sequence contains continuous or discontinuous sequences made up of one or more reference segments
What is a reference segment?	A reference segment is a contiguous section of the reference backbone within the reference sequence that corresponds to a single contiguous DNA sequence. The reference backbone is the entire reference sequence and can be composed of one or more reference segments.
What is the reference backbone?	The reference backbone is the entire reference sequence that may be comprised of many reference segments. The backbone is the first layer of the RDG and this first layer may not be modified.

Table B-5 SeqScape Manager Questions and Answers *(continued)*

Question	Answer
What does splitting the reference mean and how do I split it?	Creating a split can represent that the sequences are not contiguous; One side of the split may contain Exon3 and the other side may represent Exon8. On the ROI tab, click the base position where you would like to split the reference segment and select Split Reference Segment.
Where can I find information on the ROI tab?	You can find information by clicking **Info** on the bottom left of the ROI tab within the RDG.
Can the ROI contain negative numbering?	Yes, the ROI can contain negative numbering. You can assign a negative number to an ROI by entering the number into the ROI table of the RDG.
Can the reference sequence contain negative numbering?	The reference backbone sequence cannot contain negative numbering, however individual ROIs within the reference backbone can contain negative numbers.
How do I save GenBank files?	After finding the desired sequence at the NCBI website, select the check box to the left of the accession number. At the top of the page next to Display, select **GenBank**, then select **Send to File**. The file will be saved to the specified location. This file can then be imported into the RDG. The file can have a .gb; .fcgi; or .cgi extension.
How can I designate part of my sequence as untranslated (intronic region)?	You can designate part of the sequence as untranslated by deselecting the **Translate** check box for that ROI. You can designate a region of the sequence as untranslated by first highlighting the desired section of the sequence in ROI tab and clicking on **Add ROI**. This region will then appear in the ROI table. Select the layer where you would like this ROI to appear and deselect the **Translate** check box.
How can I change the number of the first base in the reference sequence? How can I reset the first codon?	You can designate the first base/codon in the reference segment pane of the ROI tab. This is the pane that shows selected reference sequence. Click in the box on the top left of the pane, and enter the desired number.
How can I change the translation frame?	You can change the translation frame in the ROI tab of the RDG.
Can I use an implicit reference sequence in SeqScape software?	No, SeqScape software does not support the use of an implicit reference sequence. However, you can use .ab1 files and genbank files as reference sequences.
What is a reference break?	A reference break is a break in the reference sequence between two reference segments where the reference is not contiguous.

Table B-5 SeqScape Manager Questions and Answers *(continued)*

Question	Answer
What is a translation codon table?	A table that translates amino acid and genetic codes. Refer to Appendix C, “Translation Tables.”
What is a known variant?	An AA variant or NT variant that has been previously identified in the reference.
What is a region of interest (ROI)?	An ROI is a region on the reference segment with special numbering properties used for display. The numbering for the ROI is continuous, always increases from left to right, and does not have to correspond to the numbering on the reference segment. The ROI can also contain negative numbers. ROIs can be grouped into layers for display or translational purposes
How can I configure a reference segment and ROIs within it?	After you import a reference sequence into the RDG, use the ROI tab to reconfigure a reference segment and to add ROIs.
What if I do not have variant information?	Variants are not necessary to create a reference data group. If you do import variants, they must be in a tab-delimited text file format or FASTA alignment of sequences.
File import— What kinds of files can I import into SeqScape software?	ABI sample files, tab-delimited text, and FASTA file format can be imported into the software.
File import— Can I import multiple individual text files into one specimen rather than .ab1 files?	Yes. You can import .seq format files or FASTA format files as sample files. In order to see these files in the import dialog, you need to deselect **Show .ab1 Samples File Only** box. These files can then be analyzed like normal files (except no basecalling occurs).
File import— Do my sample files need to be imported in the same reading frame as the reference sequence?	No, the imported sample files do not need to be in the same reading frame as the reference sequence
File format— Can analyzed data be used in SeqScape software?	Analyzed data can be used. However, if it is in the ABI data format (and not FASTA), any prior analysis, results, and edits will be overwritten when the files are reanalyzed using SeqScape software.

Table B-5 SeqScape Manager Questions and Answers *(continued)*

Question	Answer
File format— What is FASTA format? How can I convert non-FASTA files into the correct format?	A sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a greater-than (>) symbol in the first column. **Note:** When creating a file in Microsoft® Word, be sure to save it in text-only format (line breaks are OK, but spaces are not OK). >HumMitoCamb from 15871 to 450 (hard return) aatactcaaatgggcctgtccttgtagtataaactaatacaccagtcttgtaaaccggagatg aaaaccttttccaaggacaaatcagagaaaaagtctttaactccaccattagcacccaaag ct (hard return)
What are Analysis Settings?	The analysis settings determine the basecalling, mixed base settings, clear range, and filter settings.
What is Clear Range?	Clear range defines the range of usable sample sequence to be included in the consensus.
Can the Clear Range be modified within a project? Does changing the Clear Range require that the data be re-analyzed?	You can change the Clear Range for the entire project by applying a new Project Template (with a modified Analysis Protocol), in which case all samples will need to be re-analyzed and any sample basecall/edits will be lost. You can reset the Clear Range for an individual sample by modifying the sample's Analysis Protocol setting. In this case, the specimen containing affected sample(s) must be re-analyzed. If only the Clear Range tab was modified in the Analysis Protocol, the analysis pipeline will be started from Clear Range determination onward. So, basecalls will not be overwritten You can also change the sample Clear Range within the sample file. This will not require re-analysis. To change the clear range in this way, right-click a sample sequence and select **Set CR... at selection**, then click-drag the CR bracket to reset the Clear Range. You can also select **Tools** > **Set Clear Range** to reset the sample clear range.
What are Filter Settings?	These allow you to set the maximum percentage of mixed-bases allowed, maximum Ns allowed, minimum clear range length, and the minimum sample score for each sample. Samples failing the filter checks will not be included in the analysis.
What are Display Settings?	These control the font styles and colors for bases, electropherogram display and axis scale, display view for variants, and display views for nucleotide translation.
Can I export consensus sequences?	Yes. Select the desired consensus sequence in the Project Navigator, then select **File** > **Export**.

Library FAQs

Table B-6 Library Questions and Answers

Question	Answer
What are the requirements of the library?	The library search feature is designed for use with libraries of alleles, genotypes, or haplotypes where all the sequences in the library are variations of the same sequence. This feature is not designed for searching against a library of diverse sequences. The library must have the following characteristics: • All library sequences must be pre-aligned and equal in length. • All library sequences must be variations of the same sequence (variations must be less than 50%). • All library sequences must cover the same regions as the layer that is associated with that library (for example, for a layer that contains exons 2,3,7 and 8, a valid library should have sequences from exons 2, 3, 7 and 8. A library with sequences covering exons 2,3, 4,5, 7 and 8 would be invalid). A library will be associated to its specific layer.
Is there a minimum/maximum number of Library Search match returns that I can define?	There is no maximum, however it does not make sense to request more than the number of alleles in the library. The minimum should be 1. You can set the number of library matches to return in the Analysis Defaults > Specimen tab of the project.
What is the difference between a haploid and diploid library?	In a haploid library, all the sequences are pure base sequences. A diploid library contains both mixed base and pure base sequences. A haploid library will return 2 possible matches, while a diploid library will return 1 possible match.
Can I add a library to an open project from within the RDG properties button and see the search results instantly, or will I have to re-analyze the project for the library search to be initiated?	The library is automatically searched immediately after loading a new library for the active layer and closing the RDG dialog box.
What is the function of the Library Identification pane? How can I view the Library Identification pane?	You can use the Library Identification pane to review the crucial positions identified among the set of library matches returned against the selected specimen consensus sequence. In order to view this pane, click on a base in a specimen consensus sequence in the Project Assembly view. You can adjust the height of this pane by using the drag-split bar. The crucial position columns are hyperlinked to the specimen consensus sequence base positions highlighted by the column selector in the Project Assembly view.

Table B-6 Library Questions and Answers *(continued)*

Question	Answer
What is a crucial position error?	A crucial position error is a polymorphic position that occurs in all the allelic matches. It is the position that makes each allele unique to one another.
What is a constant position error?	A constant position error is a position in a specimen consensus sequence that is different than the corresponding position in all the allelic matches. All of the allelic matches have identical base for that particular position.

Mutation, Variant, HIM, and HFM Detection FAQs

Table B-7 Mutation, Variant, HIM, and HFM Detection Questions and Answers

Question	Answer
Does SeqScape software account for heterozygous indel mutations (HIMs)? For example, a sequence with an insertion of three bases.	Yes. SeqScape software will show possible HIM location and identity in the Analysis QC Report.
Can the SeqScape software separate the HIM sequence traces?	No. SeqScape software cannot separate HIM sequence traces. However, SeqScape software will report how many bases were deleted or inserted.

Data Analysis FAQs

Table B-8 Data Analysis Questions and Answers

Question	Answer
How do I begin analysis?	Click (Analyze) in the toolbar or select **Analysis** > **Analyze**.
What does a red line across a specimen indicate?	Strike through symbols indicate that analysis needs to be performed.
Can the SeqScape software handle gaps in sequence?	SeqScape software automatically inserts gaps in the sample and consensus sequences if these gaps are necessary to produce clean sequence alignments. Gaps should be removed before importing sequences from FASTA-formatted files.
What does the Alignment Score mean in the Analysis Report?	The alignment score shows the number of characters that were inserted in each specimen consensus to create the project alignment. A lower alignment score indicates more similarity between the specimen consensus and the reference.
How does editing affect my data? What gets updated?	If you insert, delete, or change a base within a sample, the change is reflected in the consensus sequence. All samples change to reflect the consensus edits.
How can I distinguish between edited and non-edited data?	Edited bases are displayed as lower-case letters while unedited bases are displayed in upper-case letters.
What will happen to my edited sequence when I start analysis?	Once basecalling begins, all current edits are overwritten. Changes to the analysis settings that do not require re-basecalling of the sample preserve edits and the reference sequence.
What happens if I edit a consensus base?	The base changes to lowercase in the consensus and the quality bar turns gray. All bases in the samples at that position that disagreed with the new basecall are changed to agree with the new consensus base and are shown in lowercase with a gray quality bar.
How do I remove unwanted spaces in my samples?	To remove unwanted spaces in the sample, double-click the space and press the Delete key.
What can I do if I deleted too many bases?	Repeat the analysis.
Is there an option to basecall without generating quality values?	No. All of the basecallers in SeqScape software will generate quality values. However, you do not have to display the quality values. You can choose to not display the quality values by deselecting the confidence bar icons in the Views tab of the project's Display Settings.

Table B-8 Data Analysis Questions and Answers *(continued)*

Question	Answer
Can I assemble/analyze my samples without re-basecalling my samples so that I can conserve the existing basecalls?	Yes. In order to assemble and analyze your sequences without basecalling, open the Analysis Defaults for the project and select the Specimen tab. Deselect the Basecall Samples check box.
How do I analyze samples in one project with different basecallers/dye set primer files?	To analyze samples in one project with different basecallers and dye set/primer files, you must open the Sample Manager from the Analysis menu. Select the appropriate basecaller and dye set/primer files and click **Apply**. The project must be re-analyzed for these changes to take effect. Samples and specimens with a red slash indicate an unanalyzed status. In the Sample Manager, you can also choose to Edit the Analysis Protocol for the individual samples or Apply an Analysis Protocol.
What does a red line through the specimen icon indicate?	A red line indicates that analysis has not occurred. A red line may also appear if the analysis settings have been changed and the project requires re-analysis to apply these settings. Click the green arrow (run) button at the top of the window to start analysis.
How can I edit my specimen name?	Select the specimen and select **Edit** > **Rename** or right-click and select **Edit**.
How can I delete samples or specimens?	Select the item to be deleted, then select **Edit** > **Delete**, click the Delete button on the toolbar, press the Delete key on the keyboard, or right-click and select **Delete**.
What is the TraceTuner™ basecaller module?	The ABI basecaller contains an algorithm that assigns bases and invokes a second algorithm, the TraceTuner module. TraceTuner module is responsible for generating per-base sample quality values and identifying mixed bases.
What does a red dot mean in the analyzed project?	A red dot indicates a base that has been called by the consensus caller. The consensus caller will edit this base in the relevant sample sequences of the specimen. The edited base will appear in lower case, and it will have a gray quality value bar.
Can I change the settings of the tab jump key?	You can change the settings of the tab jump key in the Views tab of the project Display Settings. You can also change the tab jump key settings when the project is open by selecting the Tab jumps to next... option on the toolbar.

Analysis Reports FAQs

Table B-9 Analysis Reports Questions and Answers

Question	Answer
How can I access my reports?	Access all reports by clicking the Report Manager button in the toolbar or by selecting **Analysis** > **Report Manager**.
What are the different reports available in SeqScape software?	SeqScape software v2.1 has the ability to generate the following reports: • Analysis QC Report • Mutations Report • AA Variants Report • Specimen Statistics Report • Sequence Confirmation Report • Base Frequency Report • Library Search Report • RDG Report • Audit Trail Report • Genotyping Report **Note:** For more information on the reports, see "Viewing the Reports" on page 7-23.
What is the Analysis QC Report?	The Analysis QC report provides a summary of the project's history. This report details the status of each specimen at each step of analysis. In addition, the Analysis QC report lists possible HIMs (heterozygous insertion/deletion mutations).
What does "Segment Score" mean in the Specimen Statistics report?	Segment Score gives an average of all the quality values within the clear range in that particular reference segment region.
What does "Coverage" mean in the Specimen Statistics and Sequence Confirmation reports?	Coverage gives a value for the number of samples in the consensus sequence.
Can I edit sequences within a project while reports are open and see the updated information in the reports instantly, or will I have to close and re-open the reports to see any changes?	The reports will stay open, and the results are updated as edits are made.

Table B-9 Analysis Reports Questions and Answers *(continued)*

Question	Answer
Why are my sample files unassembled?	If you have samples in the unassembled node of a specimen, check the Analysis QC report to determine the reason why the sample files were not assembled. The analysis QC report will show whether or not the sample assembled, as well as the reason for failure at a particular point in the analysis pipeline.

Quality Values FAQs

Table B-10 QV Questions and Answers

Question	Answer
What are quality values?	A quality value is an estimation of the certainty for a basecall in the sample (sample QV) or consensus (consensus QV).
Is there an option to basecall without generating quality values?	No, all of the basecallers in SeqScape software generate quality values. However, you can choose to not display the quality values by deselecting the confidence bar icons in the Views tab of the project's Display Settings.
How is the basecaller quality value generated?	It is derived using an algorithm that is designed to examine the certainty of basecalls. See Chapter 10, “Sample and Consensus Quality Values,” for more information.
What is the quality value equation?	$QV = -10\log_{10}(PE)$ where PE is the probability of error.
How are sample quality values generated?	They are generated using a statistical algorithm which is calibrated to estimate the certainty of basecalls.
How is a sample quality value different from the sample score?	The sample score is the average quality value of the bases in the clear range sequence for that sample. A sample quality value is a per-base estimate of basecaller accuracy.
How does the consensus quality value differ from the consensus score?	The consensus score is the average quality value of the bases in the consensus sequence for that specimen. A consensus quality value is a per-base estimate of the accuracy of the consensus-calling algorithm.

Printing and Exporting Results FAQs

Table B-11 Printing and Exporting Results Questions and Answers

Question	Answer
Printing — What can I print in SeqScape software?	You can print the views only for sample, specimen, segment, project, and complete reports. You can also print electropherograms, complete reports, and the visible data or all data for the project view.
Printing — What printers are recommended for use with SeqScape software?	An HP® 8100, 4600, 990cxi, or an Epson® 980 color printer is recommended.
File export— What can I export from SeqScape software?	User information, projects, project alignments, project templates, reports, nucleotide and amino acid variants, and libraries can be exported from the software. Refer to Appendix D, “User Privileges.”
File export— Can I export each consensus sequence individually?	Consensus sequences for a project can be exported as a group by using selecting **File** > **Export** in the Project view.
What format can I print/export reports in?	You can export reports in .pdf, .xml, .htm, or .txt file formats. You can print the exported reports or you can choose to print an open report by selecting **File** > **Print**.
Can I export and print individual .ab1 sample files from the project?	Yes. You can export and print individual .ab1 files from within the project. To do this, select the sample file in the Project Navigator view. Then click on a sample file and go to **File** > **Export** > **Sample Sequence File**. The sample file can be exported in four formats: .seq, .fsta (FASTA), .phd.1 (PHD), and .ab1. You can print individual sample files by selecting **File** > **Print**.

Audit Trail, Security, and Access Control FAQs

Table B-12 Audit Trail, Security, and Access Control Questions and Answers

Question	Answer
What security and audit trail features are included in SeqScape software?	SeqScape software v2.1 has the following security and audit trail features: • Three levels of user access • User lockout after certain time frame has passed • Password expiration • Audit trail that can be created for base change, insertion, or deletion • Audit trail also includes time/date stamp and reason for change • User name that is displayed when logged into the software • Audit Trail report
What are the access control differences as you go from Admin to Scientist to Analyst?	**Administrators** can do everything that the application possibly lets you do. This includes the admin-specific tasks: creating users, viewing and changing user details, importing and exporting users, and changing the 21CFR Authentication and Audit features. **Scientists** can do everything except the admin-specific tasks. **Analysts** can open projects and import samples, but cannot affect other master objects. For example, the analyst cannot view, modify, import, or export project templates, RDGs, analysis defaults, etc. This includes changing the RDG or analysis settings once inside a project, although you can change the basecaller and dye set/primer files from within the sample manager. The Analyst is allowed to edit the project. The Analyst also cannot perform any admin-specific tasks. For more information, see Appendix D, “User Privileges.”
Does the audit trail component add User ID and a Time/Date stamp to each entry?	The Audit Trail report does include a user ID, user first and last name, and time/date stamp for each audit event. The Audit report also includes the reason why the user modified the data, and it includes any comments that the user may have included.

Translation Tables

C

This appendix contains:

IUPAC/IUB Codes. C-2
IUPAC Diagrams. C-3
Complements. C-3
Universal Genetic Code. C-4
Amino Acid Abbreviations C-5

IUPAC/IUB Codes

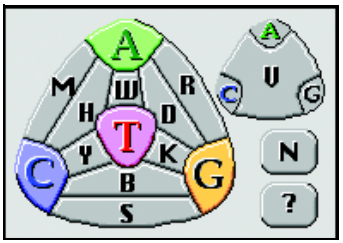
Table C-1 IUPAC/IUB Codes

Code	Translation
A	Adenosine
C	Cytidine
G	Guanosine
T	Thymidine
B	C, G, or T
D	A, G, or T
H	A, C, or T
R	A or G (puRine)
Y	C or T (pYrimidine)
K	G or T (Keto)
M	A or C (aMino)
S	G or C (Strong—3 H bonds)
W	A or T (Weak—2 H bonds)
N	aNy base
V	A, C, or G

Note: This chart is accessible from the Help menu.

IUPAC Diagrams

IUPAC



IUPAC heterozygous

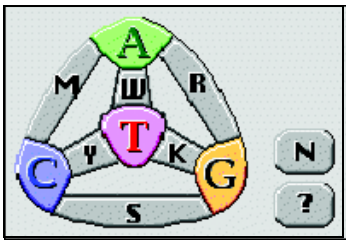


Figure C-1 IUPAC Diagrams

Complements

Table C-2 Complements

A	T	S	S
C	G	W	W
G	C		
T	A	B	V
		D	H
R	Y	H	D
Y	R	V	B
K	M	N	N
M	K		

Note: These charts are accessible from the Help menu.

Universal Genetic Code

Table C-3 Universal Genetic Codes

5' End	2nd Position				3' End
	T	C	A	G	
T	Phe	Ser	Tyr	Cys	T
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	OCH	OPA	A
	Leu	Ser	AMB	Trp	G
C	Leu	Pro	His	Arg	T
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	T
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	T
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G
Stop Codes: AMBer, OCHer, OPA					

Note: This chart is accessible from the Help menu.

Amino Acid Abbreviations

Table C-4 Amino Acid Abbreviations

Amino Acid	Three Letters	One Letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Any Amino Acid		X

Note: This chart is accessible from the Help menu.

User Privileges

D

Tables of User Privileges

This appendix contains a list of privileges for users of the three categories, Administrator, Scientist, and Analyst, when they use the ABI PRISM® SeqScape® Software.

Table D-1 Access for Admin Level

		Description of access for users of Admin level only	Admin	Scientist	Analyst
Admin only access	1	Create User Accounts	Allowed	Not Allowed	Not Allowed
	2	Exporting/Importing User Accounts			
	3	Export a Project/PT/RDG/Library from the SeqScape Manager			
	4	Import objects from outside the DataStore into the SeqScape Manger			
	5	Install SeqScape for an automated analysis system			

Table D-2 Access for Admin and Scientist Levels

		Description of access for users of Admin and Scientist levels	Admin	Scientist	Analyst
SeqScape Manager	1	Delete an object from the SeqScape manager	Allowed	Allowed	Not Allowed
	2	Delete a Project from the SeqScape manager			
	3	Save As... an object in the SeqScape manager			
	4	Create a new object in the SeqScape Manager			
	5	Create a new Project Template in SeqScape Manager			
	6	Configure analysis defaults in SeqScape manager			
	7	Deleting entries from a library in the SeqScape manager			
	8	Re-Configure an existing Project Template in the SeqScape Manager			
Analysis Protocol & Settings	9	Creating an analysis protocol	Allowed	Allowed	Not Allowed
	10	Editing an existing analysis protocol			
	11	Apply an analysis protocol to a set of samples (project/sample/specimen)			
	12	Create new Primary Seq Analysis Protocols			
	13	Set Clear range determination in Analysis settings or analysis defaults			
	14	Set Mixed Base determination in Analysis settings or analysis defaults in a Project, PT/SS Manager			

Table D-2 Access for Admin and Scientist Levels *(continued)*

		Description of access for users of Admin and Scientist levels	Admin	Scientist	Analyst
RDG	15	RDG: Import Variants and Reference into an RDG from a set of aligned FASTA files	Allowed	Allowed	Not Allowed
	16	RDG general tab: configure an RDG in general tab			
	17	RDG ROI tab: Edit a Reference Data Group (RDG): configure Layers			
	18	RDG ROI tab: Edit a Reference Data Group (RDG): configure ROIs			
	19	RDG ROI tab: Edit a Reference Data Group to use an implicit reference			
	20	RDG ROI tab: adding/modifying a Reference Segment			
	21	RDG ROI tab: Change the Reference Segment index Base in an embedded RDG			
	22	RDG ROI tab: deleting a Layer			
	23	RDG ROI tab: deleting a Reference Segment			
	24	RDG ROI tab: deleting an ROI			
	25	RDG ROI tab: Import GenBank sequences into the RDG for automated Ref Segment and feature creation			
	26	RDG NT variants Tab: Edit NT variants in an RDG			
	27	RDG NT variants Tab: Import NT variants from a Tab Delimited Text into RDG			
	28	RDG AA variants Tab: Add amino acid variants to an RDG			
	29	RDG AA variants Tab: Edit AA variants in a RDG			

Table D-2 Access for Admin and Scientist Levels *(continued)*

		Description of access for users of Admin and Scientist levels	Admin	Scientist	Analyst
RDG	30	RDG AA variants Tab: Import AA variants from a tab delimited text file into RDG	Allowed	Allowed	Not Allowed
	31	RDG variant styles tab: configure an RDG in Variant Styles tab			
Library	32	Library: overwriting/appending sequences to an existing library	Allowed	Allowed	Not Allowed
	33	Library: editing sequence data in the library			
	34	Library: exporting data from the library as a Multi-FASTA file			
	35	Library: viewing/editing library types in the Library Type manager			
	36	Library: creating a new sequence library			
Other	37	Sets General Preferences in Options	Allowed	Allowed	Not Allowed
	38	Sets Sequence Collector (Database) Preferences in Options			
	39	Add NT or AA variants from any data view			
	40	Set specimen level analysis settings			
	41	Set project level analysis settings			

Table D-3 Access for Admin, Scientist and Analyst Levels

		Description of access for users of Admin, Scientist and Analyst levels	Admin	Scientist	Analyst
Reports	1	View Reports	Allowed	Allowed	Allowed
	2	View Reports with enabled links back to primary data			
	3	View Reports while editing project			
	4	Export all reports			
	5	Export all customized reports			
	6	Print all reports			
	7	View heterozygous frame shifts links from Mutations Report			
	8	Print a report from the reports manager			
Project View/Display	9	Move sample data from one Specimen to another	Allowed	Allowed	Allowed
	10	Display SQVs and CQVs			
	11	Re-order aligned Specimen consensi			
	12	Change active Layer view			
	13	Show/hide variants that result in silent mutations			
	14	Sort Summary Table in Specimen view			
	15	Display Sample and Consensus Scores			
	16	View Amino Acid tooltips for degenerate codons			
	17	View Amino Acid Alignment in Main Window			
	18	View Library Search Results in Alignment View Identification Pane			

Table D-3 Access for Admin, Scientist and Analyst Levels *(continued)*

		Description of access for users of Admin, Scientist and Analyst levels	Admin	Scientist	Analyst
Project View/Display	19	View electropherogram data as aligned peaks	Allowed	Allowed	Allowed
	20	View all objects in Project Navigator and Main Windows			
	21	View Specimen Layout			
	22	View Specimen-Segment Assembly tab			
	23	View Unassembled data in the Project Navigator and Specimen Views			
	24	View/Navigate through electropherogram snippets			
	25	View/Navigate Specimen Segment electropherogram data			
	26	View a Project/Navigate using the Overview pane			
	27	View Samples in the Sample Manager tab			
	28	View/Navigate alignments using the display toolbar buttons			
Project-Other Controls	29	Apply a new Project Template to an existing Project	Allowed	Allowed	Allowed
	30	Create a new Project from the SeqScape Toolbar			
	31	Delete Samples in Project Navigator			
	32	Delete Specimens in Project Navigator			
	33	Export Project Alignment in FASTA format			
	34	Export Sample data in SEQ, FASTA or AB1 format			
	35	Export Specimen consensus or aligned sample sequences in FASTA format			

Table D-3 Access for Admin, Scientist and Analyst Levels *(continued)*

		Description of access for users of Admin, Scientist and Analyst levels	Admin	Scientist	Analyst
Project-Other Controls	36	Import a Text segment to a Text Specimen	Allowed	Allowed	Allowed
	37	Import Samples to Project			
	38	Import Samples to Project from Database (Sequence Collector)			
	39	Import/create a text-only Specimen			
	40	Open an embedded Settings Object inside a Project			
	41	Open an existing Project			
	42	Print wrapped nucleotide or amino acid Project Alignments			
	43	Save Project from the Menu or Toolbar			
	44	Search for text strings in any sequence data			
Editing	45	Generate an Audit Trail event	Allowed	Allowed	Allowed
	46	Project Alignment view: Change consensus basecalls			
	47	Project Alignment view: Insert or delete a space in a Reference			
	48	Project Alignment view: Insert or delete a space in a Specimen consensus			
	49	Project Alignment view: Insert/delete Consensus bases			
	50	Project Navigator: Rename Specimens			
	51	ROI tab: Rename Segments in RDG			
	52	Specimen view: Change a base in a sample			
	53	Specimen view: Change basecalls in the consensus			

Table D-3 Access for Admin, Scientist and Analyst Levels *(continued)*

		Description of access for users of Admin, Scientist and Analyst levels	Admin	Scientist	Analyst
Editing	54	Specimen view: Change the Clear Range for sample data	Allowed	Allowed	Allowed
	55	Specimen view: Insert or delete a base in a sample			
	56	Specimen view: Insert or delete bases in consensus			
	57	Undo base edits			
SeqScape Manager	58	Open the SeqScape Manager	Allowed	Allowed	Allowed
	59	Save any SeqScape Manager Object			
Library	60	View the Libraries in the SeqScape Manager	Allowed	Allowed	Allowed
	61	View results of Library search in the Project Alignment View			
Analysis Protocol and Settings	62	View the Analysis Protocol	Allowed	Allowed	Allowed
	63	Change basecaller settings in an existing Sample within a Project			
	64	Reconfigure Analysis Defaults inside a Project			
	65	Configure Display Settings in Project or SeqScape Manager			
	66	Analyze data using the BGB without basecalling samples			
	67	Analyze data using the BGB			
	68	Indicate that specific Samples are not to be basecalled			

Table D-3 Access for Admin, Scientist and Analyst Levels *(continued)*

		Description of access for users of Admin, Scientist and Analyst levels	Admin	Scientist	Analyst
Other	69	Browse/Locate data in the file system	Allowed	Allowed	Allowed
	70	Exit SeqScape			
	71	Sort items in columns in any table in SeqScape			
	72	Install SeqScape on a clean system			
	73	Upgrade SeqScape Software v1.0 or v1.1 to v2.0			
	74	Uninstall SeqScape			
	75	Launch SeqScape			
	76	Configure a sample in Data Collection for automated import into SeqScape			

Aligned Variant and FASTA File Format

E

This appendix contains:

Tab-Delimited Files E-2
FASTA File Format E-3

Tab-Delimited Files

You can import variants into the ABI PRISM® SeqScape® Software if they are in the format of a tab-delimited text file.

Creating a Variant Text File

SeqScape software tab-delimited text files must conform to the following rules:

- One variant per line
- The following tab-delimited column headings:

NT Variant Headings	AA Variant Headings
Type	Type
ROI	Layer
NT position	AA position
Reference	Reference
Variant	Variant
Style	Style
Description	Description
Used by all ROIs	

An example is provided in Figure E-1.

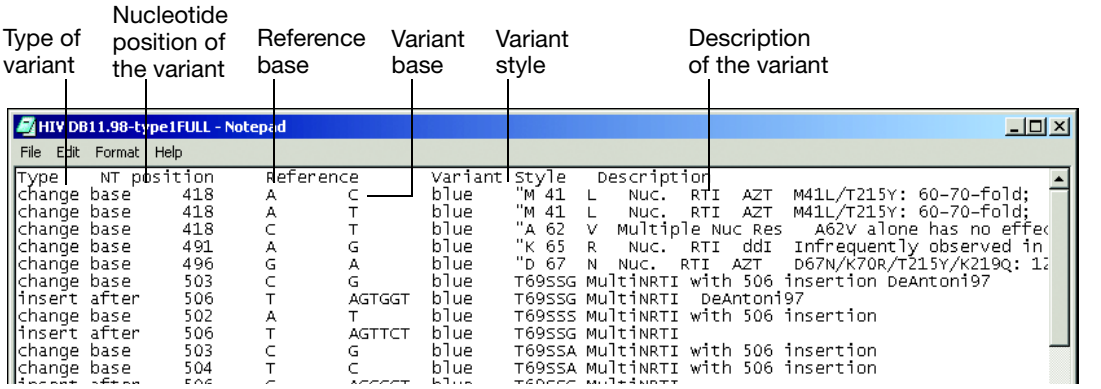

Type of variant
Nucleotide position of the variant
Reference base
Variant base
Variant style
Description of the variant
HIV DB11.98-type1FULL - Notepad
File Edit Format Help
Type NT position Reference Variant Style Description
change base 418 A C blue "M 41 L Nuc. RTI AZT M41L/T215Y: 60-70-fold;
change base 418 A T blue "M 41 L Nuc. RTI AZT M41L/T215Y: 60-70-fold;
change base 418 C T blue "A 62 V Multiple Nuc Res A62V alone has no effe
change base 491 A G blue "K 65 R Nuc. RTI ddI Infrequently observed in
change base 496 G A blue "D 67 N Nuc. RTI AZT D67N/K70R/T215Y/K219Q: 1
change base 503 C G blue T69SSG MultiNRTI with 506 insertion DeAntoni97
insert after 506 T AGTGGT blue T69SSG MultiNRTI DeAntoni97
change base 502 A T blue T69SSS MultiNRTI with 506 insertion
insert after 506 T AGTTCT blue T69SSG MultiNRTI
change base 503 C G blue T69SSA MultiNRTI with 506 insertion
change base 504 T C blue T69SSA MultiNRTI with 506 insertion

Figure E-1 Sample of Variant Tab-Delimited Text Format

FASTA File Format

Note: The information on FASTA was obtained from http://www.ncbi.nlm.nih.gov/BLAST/fasta.html

FASTA Format Description

A sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a greater-than (>) symbol in the first column.

FASTA Format Example

An example sequence in FASTA format is as follows:

>HIV HXB2 Prt-RT1(1-320)
cctcaggtcactctttggcaacgacccctcgtcacaataaagataggggggcaactaaaggaag
ctctattagatacaggagcagatgatacagtattagaagaaatgagtttgccaggaagatggaaa
ccaaaaatgatagggggaattggaggttttatcaaagtaagacagtatgatcagatactcatagaa
atctgtggacataaagctataggtacagtattagtaggacctacacctgtcaacataattggaaga
aatctgttgactcagattggttgcactttaaattttcccattagccctattgagactgtaccagtaaaat
taaagccaggaatggatggcccaaaagttaaacaatggccattgacagaagaaaaaataaaag
cattagtagaaatttgtacagagatggaaaaggaagggaaaatttcaaaaattgggcctgaaaat
ccatacaatactccagtatttgccataaagaaaaaagacagtactaaatggagaaaattagtagatt
tcagagaacttaataagagaactcaagacttctgggaagttcaattaggaataccacatcccgca
gggttaaaaaagaaaaaatcagtaacagtactggatgtgggtgatgcatatttttcagttcccttag
atgaagacttcaggaagtatactgcatttaccatacctagtataaacaatgagacaccagggatta
gatatcagtacaatgtgcttccacagggatggaaaggatcaccagcaatattccaaagtagcatg
acaaaaatcttagagccttttagaaaacaaaatccagacatagttatctatcaatacatggatgattt
gtatgtaggatctgacttagaaatagggcagcatagaacaaaaatagaggagctgagacaacat
ctgttgaggtggggacttaccacaccagacaaaaaacatcagaaagaacctccattccctttggat
gggttatgaactccatcctgataaatggacagtacagcctatagtgctgccagaaaaagacagct
ggactgtcaatgacatacagaagttagtggggaaattgaattgggcaagtcagatttacccaggg
attaaagtaaggcaattatgtaaactccttagaggaaccaaagcactaacagaagtaataccacta
acagaagaagcagagctagaactggcagaaaacagagagattctaaaagaaccagtacatgg
agtgtattatga

IUB/IUPAC Codes

Sequences are expected to be represented in the standard IUB/IUPAC amino acid and nucleic acid codes, with the following exceptions:

- Lower-case letters are accepted and are mapped into uppercase
- In amino acid sequences, U and * (asterisk) are acceptable letters (see below)

Note: Although FASTA codes allow a hyphen or dash to represent a gap in nucleotide sequences, this practice is not acceptable for using FASTA format in SeqScape software.

Before importing a sequence, any numerical digits or spaces in the sequence need to be either removed or replaced by appropriate letter codes (for example, N for unknown nucleic acid residue or X for unknown amino acid residue).

Supported Nucleic Acid Codes

Table E-1 Accepted Nucleic Acid Codes:

Character …	Codes for …
A	Adenosine
C	Cytidine
G	Guanine
T	Thymidine
U	Uridine
R	GA (purine)
Y	TC (pyrimidine)
K	GT (keto)
M	AC (amino)
S	GC (strong)
W	AT (weak)
B	GTC
D	GAT

Table E-1 Accepted Nucleic Acid Codes: *(continued)*

Character …	Codes for …
H	ACT
V	GCA
N	AGCT

Supported Amino Acid Codes

Table E-2 Accepted Amino Acid Codes:

Character …	Codes for …
A	Alanine
B	Aspartate or asparagine
C	Cystine
D	Aspartate
E	Glutamate
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine

Table E-2 Accepted Amino Acid Codes: *(continued)*

Character …	Codes for …
T	Threonine
U	Selenocysteine
V	Valine
W	Tryptophan
Y	Tyrosine
Z	Glutamate or glutamine
X	Any
*	Translation stop
-	Gap of indeterminate length

Creating a Multi-Aligned FASTA File

F

This appendix contains:

Introduction F-1
Method 1: Create a Multi-Aligned FASTA File Using SeqScape Software F-2
Method 2: Create a Multi-Aligned FASTA File Using ClustalX Software F-4

Introduction

To use the Library Search feature in ABI PRISM® SeqScape® software, you will need to create a sequence library. All the sequences for the library must have equal lengths and must be aligned.

If you have sequences of equal length and they are already aligned, go directly to SeqScape software to create your sequence library.

This appendix contains the instructions for two different methods of preparing your sequences before creating a library.

- **Method 1** – Using SeqScape software to create a multi-aligned FASTA file
- **Method 2** – Using ClustalX software to create a multi-aligned FASTA file

About ClustalX Software

ClustalX software is a powerful multiple sequence alignment program available free of charge via the Internet. See: Jeanmougin *et al*. (1998) *Trends Biochem. Sci.* **23**, 403-5.

Method 1: Create a Multi-Aligned FASTA File Using SeqScape Software

1. Create a Multi-Aligned FASTA File

To create a multi-aligned FASTA file using SeqScape software:

Note: You can try this procedure by obtaining HLA sequences from GenBank database. The file names are listed below as bullet-point items.

1. Obtain the reference sequence:
   - gi|512471|emb|X76776.1|HSHLADMBG
2. Obtain files that contain the sequences that will be used in your sequence library (.txt or .fsta files):
   - gi-1045472-gb-U32663.1-HSU32663-Human-MHC-class-II-antigen-HLA-DM-beta-chain-(HLA-DMB)-gene
   - gi-1373022-gb-U31743.1-HSU31743-Human-HLA-DMB-variant-gene
   - gi-2315188-emb-Y14395.1-HSHLADMB3-Homo-sapiens-HLA-DMB-gene
   - gi-512471-emb-X76776.1-HSHLADMBG-H.sapiens-HLA-DMB-gene
   - gi-881918-gb-U16762.1-HSU16762-Human-(DMB)-gene
3. Launch SeqScape software and create an RDG using the file obtained in step 1.
4. Create a project template.
5. Start a new project using the project template created.
6. Add the files obtained in step 2 and analyze.
7. Select **File** > **Export**.
8. Select **Project Alignment-Nucleotides**.

You can now use this exported file to create a library in SeqScape software.

2. Create Your Library in SeqScape Software

To create your library in SeqScape software:

1. Launch SeqScape software.
2. Select **Tools** > **SeqScape Manager**.
3. Select the **Library** tab and click **New**.
4. Enter a library name.
5. Select the **Entries** tab and click **Import**.
6. Select **Display All File Types**, then navigate to and open the multi-aligned FASTA file.

The software creates the library and displays the number of sequences added.

Method 2: Create a Multi-Aligned FASTA File Using ClustalX Software

1. Obtain ClustalX Software

To obtain ClustalX software:

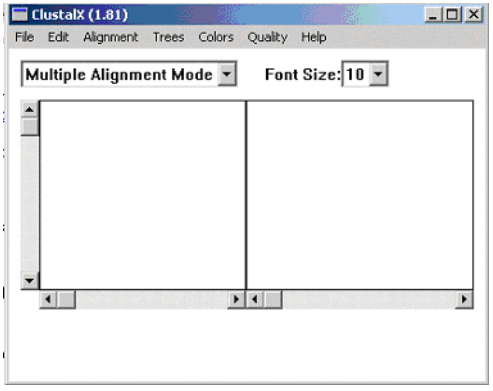
1. Launch an internet browser and search for "ClustalX" or go to: http://inn-prot.weizmann.ac.il/software/
2. Select ClustalX software from the list of software products available.
3. Select the ClustalX software for your operating system.
4. Download the software to your computer.

2. Create a Multi-Aligned FASTA File

To create a multi-aligned FASTA file using ClustalX software:

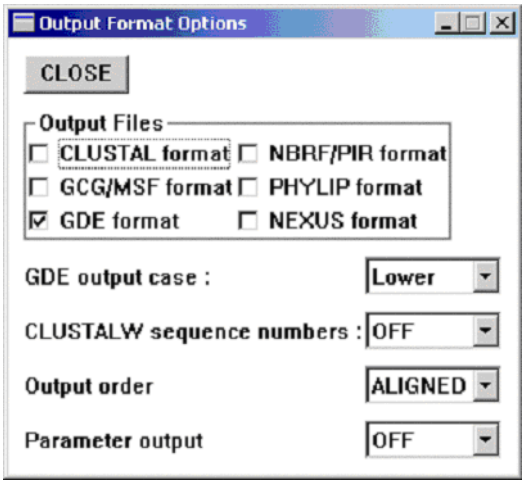
Note: You can try this procedure by obtaining HLA sequences from ftp://ftp.ebi.ac.uk/pub/databases/imgt/mhc/hla/ and selecting "DRB_nuc.fasta."

1. Open the folder that you downloaded.
2. Double click the ClustalX icon.

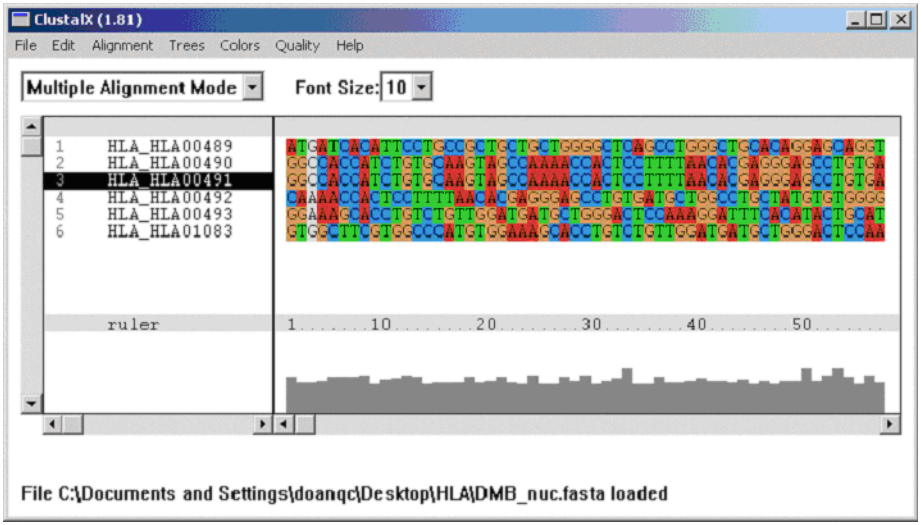


3. Select **Multiple Alignment Mode**.
4. Select **Alignment** > **Output Format Options**.

5. Select **GDE format** and set the other parameters as indicated below.

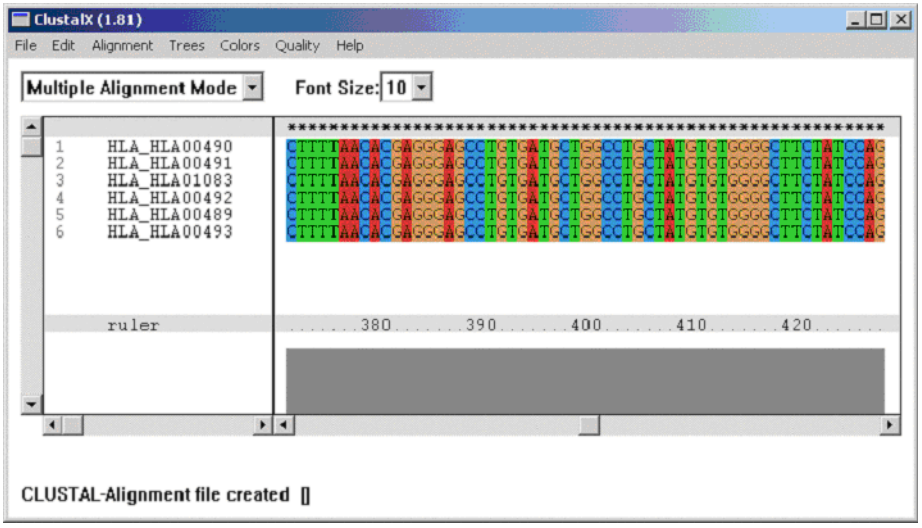


6. Select **File** > **Load Sequences**. To obtain a file, go to http://www.ebi.ac.uk/imgt/hla/ then select the **Download** tab.
7. Browse to the location of your multiple sequence file and select the file.



8. Select **Edit** > **Select All Sequences**.

9. Select **Alignment** > **Alignment Parameters** > **Reset All Gaps Before Alignment**.
10. Select **Alignment** > **Do Complete Alignment**.
11. A dialog box appears asking you to select the location where the results should be saved. Select the location.
12. Click the **ALIGN** button and wait for the alignment to be completed. Notice that dashes have been introduced where bases were missing.



13. Navigate to the file that was created. The extension on this file is .gde.

3. Edit Characters in the .gde File

To edit characters in the .gde file from ClustalX using a text editor:

1. Launch a text editor, such as Microsoft® Word.
2. Open the .gde file that you just created.
3. Search for all "#" and replace with ">".
4. Save this file as a text document with the file extension .txt or .fsta.

Note: The format of the sequences in this document is commonly referred to as a multi-aligned fasta format.

You can now use this file to create a library in SeqScape software.

4. Create Your Library in SeqScape Software

To create your library in SeqScape software:

1. Launch SeqScape software.
2. Select **Tools** > **SeqScape Manager**.
3. Select the **Library** tab and click **New**.
4. Enter a library name.
5. Select the **Entries** tab and click **Import**.
6. Select **Display All File Types**, then navigate to and open the multi-aligned FASTA file.

The software creates the library and displays the number of sequences added.

Software Warranty

G

Computer Configuration

Applied Biosystems supplies or recommends certain configurations of computer hardware, software, and peripherals for use with its instrumentation. Applied Biosystems reserves the right to decline support for or impose extra charges for supporting nonstandard computer configurations or components that have not been supplied or recommended by Applied Biosystems. Applied Biosystems also reserves the right to require that computer hardware and software be restored to the standard configuration prior to providing service or technical support. For systems that have built-in computers or processing units, installing unauthorized hardware or software may void the Warranty or Service Plan.

Limited Product Warranty

Limited Warranty

Applied Biosystems warrants that for a period of ninety (90) days from the date the warranty period begins, its ABI PRISM® SeqScape® Software Version 2.1 will perform substantially in accordance with the functions and features described in its accompanying documentation when properly installed on the instrument system for which it is designated, and that for a period of ninety (90) days from the date the warranty period begins, the tapes, diskettes, or other media bearing the software product will be free of defects in materials and workmanship under normal use. If buyer believes that it has discovered a failure of the software to satisfy the foregoing warranty, and if buyer notifies Applied Biosystems of such failure in writing during the ninety (90) day warranty period, and if Applied Biosystems is able to reliably reproduce such failure, then Applied Biosystems, at its sole option, will either (i) provide any software corrections or "bug-fixes" of the identified failure, if and when they become commercially available, to buyer free of charge, or (ii) notify buyer that Applied Biosystems will accept a return of the software from the buyer and, upon such return and removal of the

software from buyer's systems, terminate the license to use the software and refund the buyer's purchase price for the software. If there is a defect in the media covered by the above warranty and the media is returned to Applied Biosystems within the ninety (90) day warranty period, Applied Biosystems will replace the defective media. Applied Biosystems does not warrant that the software will meet buyer's requirements or conform exactly to its documentation, or that operation of the software will be uninterrupted or error free.

Warranty Period Effective Date

Any applicable warranty period under these sections begins on the earlier of the date of installation or ninety (90) days from the date of shipment for software installed by Applied Biosystems personnel. For all software installed by the buyer or anyone other than Applied Biosystems, the applicable warranty period begins the date the software is delivered to the buyer.

Warranty Claims

Warranty claims must be made within the applicable warranty period.

Warranty Exceptions

The above warranties do not apply to defects resulting from misuse, neglect, or accident, including without limitation: operation outside of the environmental or use specifications, or not in conformance with the instructions for the instrument system, software, or accessories; improper or inadequate maintenance by the user; installation of software or interfacing, or use in combination with software or products, not supplied or authorized by Applied Biosystems; and modification or repair of the product not authorized by Applied Biosystems.

THE FOREGOING PROVISIONS SET FORTH APPLIED BIOSYSTEMS' SOLE AND EXCLUSIVE REPRESENTATIONS, WARRANTIES, AND OBLIGATIONS WITH RESPECT TO ITS PRODUCTS, AND APPLIED BIOSYSTEMS MAKES NO OTHER WARRANTY OF ANY KIND WHATSOEVER, EXPRESSED OR IMPLIED, INCLUDING WITHOUT LIMITATION, WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHETHER ARISING FROM A STATUTE OR OTHERWISE IN LAW OR FROM A COURSE OF DEALING OR USAGE OF TRADE, ALL OF WHICH ARE EXPRESSLY DISCLAIMED.

Warranty Limitations

THE REMEDIES PROVIDED HEREIN ARE THE BUYER'S SOLE AND EXCLUSIVE REMEDIES. WITHOUT LIMITING THE GENERALITY OF THE FOREGOING, IN NO EVENT SHALL APPLIED BIOSYSTEMS BE LIABLE, WHETHER IN CONTRACT, TORT, WARRANTY, OR UNDER ANY STATUTE (INCLUDING WITHOUT LIMITATION ANY TRADE PRACTICE, UNFAIR COMPETITION, OR OTHER STATUTE OF SIMILAR IMPORT) OR ON ANY OTHER BASIS, FOR DIRECT, INDIRECT, PUNITIVE, INCIDENTAL, MULTIPLE, CONSEQUENTIAL, OR SPECIAL DAMAGES SUSTAINED BY THE BUYER OR ANY OTHER PERSON OR ENTITY, WHETHER OR NOT FORESEEABLE AND WHETHER OR NOT APPLIED BIOSYSTEMS IS ADVISED OF THE POSSIBILITY OF SUCH DAMAGES, INCLUDING WITHOUT LIMITATION, DAMAGES ARISING FROM OR RELATED TO LOSS OF USE, LOSS OF DATA, FAILURE OR INTERRUPTION IN THE OPERATION OF ANY EQUIPMENT OR SOFTWARE, DELAY IN REPAIR OR REPLACEMENT, OR FOR LOSS OF REVENUE OR PROFITS, LOSS OF GOOD WILL, LOSS OF BUSINESS, OR OTHER FINANCIAL LOSS OR PERSONAL INJURY OR PROPERTY DAMAGE.

NO AGENT, EMPLOYEE, OR REPRESENTATIVE OF APPLIED BIOSYSTEMS HAS ANY AUTHORITY TO MODIFY THE TERMS OF THIS LIMITED WARRANTY STATEMENT OR TO BIND APPLIED BIOSYSTEMS TO ANY AFFIRMATION, REPRESENTATION, OR WARRANTY CONCERNING THE PRODUCT THAT IS NOT CONTAINED IN THIS LIMITED WARRANTY STATEMENT, AND ANY SUCH MODIFICATION, AFFIRMATION, REPRESENTATION, OR WARRANTY MADE BY ANY AGENT, EMPLOYEE, OR REPRESENTATIVE OF APPLIED BIOSYSTEMS WILL NOT BE BINDING ON APPLIED BIOSYSTEMS UNLESS IN A WRITING SIGNED BY AN EXECUTIVE OFFICER OF APPLIED BIOSYSTEMS.

THIS WARRANTY IS LIMITED TO THE BUYER OF THE PRODUCT FROM APPLIED BIOSYSTEMS AND IS NOT TRANSFERABLE.

Some countries or jurisdictions limit the scope of or preclude limitations or exclusion of warranties, of liability, such as liability for gross negligence or willful misconduct, or of remedies or damages, as or to the extent set forth above. In such countries and jurisdictions, the limitation or exclusion of warranties, liability, remedies or damages set forth above shall apply to the fullest extent permitted by law, and shall not apply to the extent prohibited by law.

Glossary

ABI basecaller	An algorithm used in earlier versions of ABI PRISM® DNA Sequencing Analysis and ABI PRISM® SeqScape® Software.
administration	The functions of SeqScape software relating to installing, removing, or updating the application.
aligned allele library	A collection of aligned sequences that are all variations of the same sequence. This is the only type of library supported in SeqScape v2.0. An aligned allele library differs from a library of diverse sequences such as a library of different gene sequences, and is also different from a library of unaligned sequences.
alignment	The aligned reference sequence together with the aligned specimen consensus sequences.
alignment display	A table of IUB codes, space characters, blanks, and dots showing how the sequences within a project are aligned.
alignment score	The number of mismatches between the aligned reference and the aligned consensus sequence for a given specimen.
allele	An alternative form of a genetic locus.
analysis	The complete procedure that SeqScape Software performs in a batch-wise manner on sample data.
analysis defaults	The default analysis settings that are stored in a project template.
analysis protocol	The default settings (basecalling, mixed base identification, clear range and trimming, and filtering) that govern sample analysis.
analysis settings	The parameters that govern the basecalling, trimming, filtering, and assembly of the analysis.

assembly	The set of aligned and overlapping sample data that result from the sequencing of one PCR product or clone.
Assembly view	Shows the specimen consensus sequence as well as the aligned sample sequences. Electropherograms and quality values can also be viewed.
basecaller	An algorithm that determines the bases within a sequence during analysis. There are two types of basecallers: KB basecallers and ABI basecallers.
clear range	The region of sequence that remains after excluding the low-quality or error-prone sequence at both the 5' and 3' ends.
comparison	The relationship between the aligned specimen consensi and the reference sequence and the associated reference data.
consensus quality values	See quality values.
consensus caller	The analysis algorithm that is responsible for generating an accurate consensus sequence with per-base quality values.
consensus sequence	The output of the assembly from a biologically related group of samples.
constant position	A position in the library alignment that is identical for every allele in the library. See polymorphic position.
constant position error	A position in a specimen consensus sequence that corresponds to a constant position in the library and that disagrees with the library at that position.
contig	The set of aligned and overlapping sample data that results from the sequencing of one PCR product or clone. Also known as an assembly.
crucial position	A position in a specimen consensus sequence that differs among the set of matches returned after a library search.
display settings	The parameters that govern the display of the data and results.

DyeSet/Primer file	The DyeSet/Primer file compensates for the mobility differences between the dyes and primers and corrects the color-code changes due to the chemistry used to label the DNA. DyeSet/Primer files are sometimes referred to as mobility files.
export	Moving the data or settings from inside the SeqScape Software Data Store to outside the SeqScape Software Data Store either in .ctf or .txt format.
FASTA format	A standard text-based file format for storing one or more sequences.
filtered sample sequence	A sample that has been processed by the basecaller/factura/filter algorithms of the pipeline.
genotype library	A library where the allele sequences are either pure-base or mixed-base sequences. When searching against a genotype library, SeqScape attempts to find the best matches to the consensus sequence without trying different allele combinations. **Note:** This term is not used by SeqScape software.
haplotype library	A library where the allele sequences are completely pure-base sequences. When searching against a haplotype library, SeqScape attempts to combine haplotypes two at a time to find the best genotype match to the consensus sequence.
IUB/IUPAC	International Union of Biochemistry/International Union of Pure and Applied Biochemistry. More information can be found at http://www.chem.qmw.ac.uk/iubmb/misc/naseq.html#300.
KB basecaller	A new algorithm that calculates mixed or pure bases and determines sample quality values.
layout view	Shows the layout of the sample assembly with arrows indicating the placement and orientation of samples.
library match	The name of one allele or the combination of two alleles (depending on the library type) that agree closely with the specimen consensus sequence.

nibbler	The algorithm that sets the clear range for each sample using the clear range settings specified in the analysis settings.
polymorphic position	A position in the library alignment that differs for some alleles in the library. See constant position.
project	A group of related sequences that share the same reference or for which there is no explicit reference.
project summary sequence	A summary of the alignment of the specimen consensi.
project template	Contains an RDG, analysis defaults, display settings, and output settings.
quality values	Measure of certainty of the basecalling and consensus-calling algorithms. Higher values correspond to lower chance of algorithm error. Sample quality values refer to the per-base quality values for a sample, and consensus quality values are per-consensus quality values.
reference	A nucleotide string that has the following attributes: it is contiguous, it is not editable, its orientation determines the project orientation, and it is stored in the RDG.
reference associated data	The things that are related or assigned to a particular base or ranges of bases on a reference. There are two types of reference associated data: structural and variant.
Reference Data Group (RDG)	The data that contains the reference and the reference associated data.
Report Manager	A window that contains nine separate reports detailing the success or failure of various portions of the analysis, statistics, mutations, AA variants, and library search information.
sample data	The output of a single lane or capillary on a sequencing instrument that will be input into SeqScape Software.

Sample Manager	A window that displays sample file name, name and specimen; last used basecaller and DyeSet/Primer files; calculated basecalling results (spacing, peak 1, start and stop); and assembly status. The sample name, basecaller, and/or DyeSet/Primer file can be changed here.
sample quality values	See quality values.
sample score	The average of the per-base quality values for the bases in the clear range sequence for the sample.
sample view	A view in the SeqScape software where you can see attributes of each AB1 file including its annotation, sequence, features, raw data, and electropherogram data.
segment	A contiguous segment of the reference sequence corresponding to a single contiguous DNA sequence.
SeqScape Manager	The software component that manages the following settings: SeqScape Software projects, project template, RDG, analysis defaults, and display settings.
space character	A character in an aligned sequence is either an IUB code or space, perhaps shown as a dash (-). A space indicates a deleted base in this string or, equivalently, an inserted base in one of the other aligned strings.
specimen	The container that holds all the sample data as assembled contigs from a biological source or PCR product.
specimen (consensus) quality value	See quality values.
specimen (consensus) score	The average overall of the consensus quality values in the consensus sequence.
specimen consensus sequence	The output of the consensus-calling algorithm from a biologically related group of samples.
specimen report	A concatenated list of all the reported information on a per specimen basis.

specimen view	A view in SeqScape software where you can see the consensus sequence and all sample files that were used to create that consensus sequence.
summary sequence	The summary consensus sequence for the entire library alignment. Pure positions in the summary sequence correspond to *constant positions* and mixed-base positions in the summary sequence correspond to *polymorphic positions*.

Index

Numerics

310 Basecaller and DyeSet/Primer files A-5
3100 Basecaller and DyeSet/Primer files A-10
3100/3100-Avant software integration 11-2
3100-Avant Basecaller and DyeSet/Primer files A-13
3700 Basecaller and DyeSet/Primer files A-15
3730 automation required files
  analysis protocol 11-4
  instrument protocol 11-4
  results group 11-4
3730/3730xl Basecaller and DyeSet/Primer files A-17
3730/3730xl software integration 11-2
377 Basecaller and DyeSet/Primer files A-8

A

AA variants
  entering 4-36
  importing 4-38
AA Variants Report 7-31
ABI basecaller 1-4, 10-3
  defined A-2, Glossary-1
ABI data files, importing sample data 6-15
adding
  samples 8-6
  specimens 6-11
  variant, in project 6-31
adjusting clear range 8-16
  with the mouse 8-18
administration, defined Glossary-1
administrator privileges 4-12, D-1
aligned
  allele library, defined Glossary-1
  FASTA files, using in library 4-20
  variants E-2 to E-4
alignment
  defined Glossary-1
  display, defined Glossary-1
  score, defined Glossary-1
allele, defined Glossary-1
amino acid abbreviations C-5
analysis
  cumulative QV scoring in reports 10-11
  defined Glossary-1
  display workflow 3-2
  protocol, defined Glossary-1
  settings, defined Glossary-1
  settings, specifying 3-12
  viewing reports 7-23
  what it entails 6-3
analysis and report questions B-18
analysis defaults
  creating new 3-3
  defined Glossary-1
  individual samples 3-17
  setting 3-14
analysis parameters
  about 8-3
  changing 8-7
analysis protocol
  applying 8-12
  creating 3-3
  editing 8-7
Analysis Protocol Editor tabs 3-4
Analysis QC Report 7-26, 10-11
analysis, changing
  base call settings 3-4
  clear range method 3-8
  filter settings 3-10
  mixed base settings 3-8

Analyst privileges D-5
existing users 2-8
analyzing
data 6-21
project 6-1
Annotation tab, sample 6-12
applying
analysis protocol 8-12
template to existing project 6-23
Assembly view, defined Glossary-2
assembly, defined Glossary-2
assigning styles to variants 4-39
At PCR Stop check box 3-5
audit trail 1-3
Audit Trail Report 7-38
Authentication & Audit, setting 2-15
autoanalysis 11-2
Autoanalysis Manager
components 11-24
explained 11-2
automated analysis 11-2
automatic analysis, before you start 11-4

B

Base Frequency Report 7-34
basecaller
ABI 1-4
ABI, defined A-2
and DyeSet/Primer, compatibility A-4
defined Glossary-2
KB 1-4
KB, defined A-2
Basecalling tab, described 3-4
before creating new RDG, requirements 4-12

C

changing
basecaller files 8-6
DyeSet/Primer files 8-6
settings within project, examples 5-6
user information 2-19
Clear Range
tab, described 3-4
widget, using 8-17
clear range
adjusting 8-16
changing 8-17
defined Glossary-2
ClustalX software, using F-1
comparison, defined Glossary-2
compatibility, basecaller and DyeSet/Primer A-4
complements for reference C-3
completing, plate record 11-20
components of project 6-4
computer
configuration requirement G-1
technical support for altered configuration G-1
consensus
caller, defined Glossary-2
quality values, defined Glossary-2
score 10-5
sequence, defined Glossary-2
consensus quality values
consensus score 10-5
explained 10-5
consensus sequence
editing in project view 8-15
editing in specimen view 8-14
importing assembled sequences 6-20
importing text 6-20
constant
position error, defined Glossary-2
position, defined Glossary-2
contig, defined Glossary-2

creating
a library 4-20
an RDG, about 4-34
analysis protocol for autoanalysis 11-11
analysis protocols 3-3
instrument protocol for autoanalysis 11-9
new layers 4-24
new NT variants 4-30
new project using project template 6-10
new users 2-13
plate record for autoanalysis 11-18
project 6-1
project template 5-1
project, before you begin 6-3
project, using New Project Wizard 6-5
RDG, administrator privileges 4-12
RDG, scientist privileges 4-12
results group for autoanalysis 11-14
crucial position, defined Glossary-2
customizing
data 7-44
header/footer, reports 9-11
reports 7-42

D

data
analyzing 6-21
display conventions 7-3
editing 8-1
reanalyzing 8-1
saving 8-21
sources, resequencing projects 1-6
when to edit 8-13
data collection software
integration with SeqScape 11-2
more information 11-8
default directory, setting up 2-20
defining an ROI 4-15
deleting
a layer 4-16
an ROI 4-16
reference segment 4-16
display settings
defined Glossary-2
specifying 3-18
displaying
sample views 7-15
segment views 7-11
Dye Primer chemistry, files A-12, A-16
Dye Terminator chemistry, files A-5, A-13
DyeSet/Primer files
about parameter 8-7
defined Glossary-3
naming conventions A-2

DyeSet/Primer files for specific instruments A-1

E

editing
analysis protocol 8-7
data 8-1
data, workflow 8-2
existing users, privileges 2-8
expanded display, viewing 7-11
export, defined Glossary-3
exporting
about 6-34
all reports automatically 9-10
data file name and format options 9-3
from SeqScape Manager 6-34
project alignment 9-4
projects 9-4
report file name and format options 9-8
reports 9-9
samples 9-7
segments 9-6
specimens 9-5
extended reference data group 1-3
Extension Penalties, described 3-12

F

FASTA
codes E-4
creating a multi-aligned FASTA file F-1
file formats E-2 to E-4
format description E-3
format, defined Glossary-3
supported amino acid codes E-5 to E-6
supported nucleic acid codes E-4
text 4-34
features
new 1-2
file sharing, data collection and SeqScape 11-28
filter settings, table of 3-11
Filter tab, described 3-4
filtered sample sequence, defined Glossary-3
first time user 2-10
format
FASTA example E-3
options, reports 9-8
frequently asked questions B-1

G

Gap and Extension Penalties, described 3-12
GenBank
downloading file 4-5
features 4-4
general questions B-5
General tab, described 3-4
genetic analyzer applications 1-5
Genotyping Report 7-39

H

haplotype library, defined Glossary-3
hard drive partitions 2-4
hardware and software requirements 2-3

I

importing
AA variant 4-38
and exporting, about 6-34
from SeqScape Manager 6-34
NT variant 4-32
reference segment 4-13
samples automatically 6-11
variants 6-28
installation preparation 2-6
installing
first time 2-7
preparation for 2-6
SeqScape software 2-5
upgrade 2-8
instrument protocol, creating for autoanalysis 11-9
integrating SeqScape and data collection software 11-2
integration automation 1-3
invalid characters in names 2-10
IUB/IUPAC, defined Glossary-3
IUPAC diagrams C-3
IUPAC/IUB codes C-2

K

KB basecaller 1-4, 10-3
defined A-2, Glossary-3
key codes
amino acid abbreviations C-5
complements C-3
IUPAC diagrams C-3
IUPAC/IUB codes C-2
translation tables C-1
Universal Genetic Code C-4

L

launching the software 2-10
layer
creating 4-24
deleting 4-16
pane, descriptions 4-18
layout view, defined Glossary-3

learning software, wizard 4-6
library
about 4-20
creating 4-20
linking 1-7
match, defined Glossary-3
searching 1-3
setting up 4-21
Library Search Report 7-35
license and warranty, rights and responsibilities 2-2
login process, user 2-10

M

main toolbar 2-24
menu structure 2-26
Mixed Bases
settings, specifying 3-7
tab, described 3-4
mixed bases, identification 8-9
mobility files, selecting A-13 to A-17
Mutations Report 7-28, 10-12

N

new
AA variants, entering 4-36
features 1-2
layers, creating 4-24
NT variants, creating 4-30
project template, about 5-3
New Project Wizard, creating project 6-5
new users
creating 2-13
logging in 2-21
nibbler, defined Glossary-4
NT variants
about 4-29
creating 4-30

P

password protection 1-3
performing analysis 1-7
plate record
completing 11-20
creating for autoanalysis 11-18
polymorphic position, defined Glossary-4
prepare for installation 2-6
print preview 9-16
printing
data views from a project 9-14
reports 9-16
privileges
administrator D-1
analyst D-5
scientist D-2
using SeqScape D-1
project
components 6-4
defined Glossary-4
overview, workflow 6-2
summary sequence, defined Glossary-4
project template
about 5-3, 6-10
creating 5-1
creating new analysis defaults 3-3
creating new project 6-10
creating, procedure 5-4
defined Glossary-4
importing and exporting 6-34
saving 5-5
Project window, overview 2-23
projects
adding specimens 6-11 to 6-12
adding variants 6-29
creating specimens 6-12
displaying views 7-4
expanded display, viewing 7-11
exporting 9-4
importing variants 6-32 to 6-33
printing data views 9-14
reanalyzing with different template 6-22
saving your data 7-21
viewing results 7-25

Q

quality values (QV) 10-1 to 10-13
consensus quality values 10-5

cumulative QV scoring in reports 10-11 to 10-13
customizing display bars 10-7
defined Glossary-4
displaying QVs 7-3
sample quality values 10-3
table of values 10-3
questions
analysis B-18
frequently asked B-1
general B-5
reports B-18
SeqScape Manager B-9

R

RDG (Reference Data Group)
about 4-3
creating 4-1, 4-5
creating new 4-5
new, using SeqScape Manager 4-12
new, using the wizard 4-6
saving copy 4-41
Wizard 4-5
RDG Properties dialog box
ROI tab, graphic 4-14
RDG Report 7-37
reanalyzing
data 8-1
data, workflow 8-2
renaming and saving project 6-22
reference
defined Glossary-4
sequence, described 4-12
reference associated data, defined Glossary-4
reference break, adding in sequence 4-27
Reference Data Group (RDG)
adding variants 6-29
defined Glossary-4
incorporating variants into projects 6-25
reference segment
deleting 4-16
importing 4-13
pasting 4-15
setting up 4-10
registering software, recording number 2-2
removing software 2-9
Report Manager, defined Glossary-4
reports
AA Variants 7-31
Analysis QC 7-26, 10-11
Audit Trail 7-38
Base Frequency 7-34
customizing 7-42
customizing header/footer 9-11
exporting 9-9
exporting all reports automatically 9-10
file name and format options 9-8
format options 9-8
Genotyping 7-39
Library Search 7-35
Mutations 7-28, 10-12
new features 1-3
print preview 9-16
printing 9-16
RDG 7-37
Sequence Confirmation 7-33
Specimen Statistics 7-32, 10-13
types 7-22
viewing 7-23
viewing the results 7-23 to 7-39
requirements, hardware and software 2-3
resequencing
applications, common 1-5
data 1-5
Results Group Editor, completing for autoanalysis 11-14
ROI
defining 4-15
deleting 4-16
tab, graphic 4-14
ROI pane
columns, described 4-19
descriptions 4-18
running an analysis 6-21

S

sample
- bases, editing 8-14
- data, defined Glossary-4
- editing 8-14 to 8-21
- exporting 9-7
- IDs 6-11
- importing automatically 6-11
- names 6-11
- quality values, defined Glossary-5
- results, viewing annotation results 7-15
- score 10-3
- score, defined Glossary-5
- view, defined Glossary-5

Sample Manager
- defined Glossary-5
- viewing analysis parameters 8-4

sample quality values
- explained 10-3
- sample score 10-3

Save To Manager As button, using 4-42

saving
- a copy of RDG 4-41
- data 8-21
- project template within project 5-5
- RDG 4-41
- template 5-5

scheduling a run 11-6

scientist privileges 4-12, D-2

segment
- defined Glossary-5
- exporting 9-6
- views, displaying 7-11

SeqScape
- menus 2-26
- typical workflow 2-28

SeqScape Manager
- creating new analysis defaults 3-3
- creating new project templates 5-4
- creating RDG 4-5, 4-12
- creating reference using aligned sequences 4-34 to 4-36
- defined Glossary-5
- exporting from 6-34
- importing to 6-34
- questions B-9
- RDG (Reference Data Group) Wizard 4-5
- window, described 2-22

Sequence Confirmation Report 7-33

sequence editing 8-13

sequencing
- data, automating the analysis 11-2
- mobility files A-13 to A-17

Set Clear Range, using 8-19

setting analysis defaults 3-14

setting up
- default directory 2-20
- library 4-21
- new project, using New Project Wizard 6-5

software
- integration 11-2
- overview, structure 2-23
- registering 2-2
- relationships for autoanalysis 11-3
- upgrading 2-8
- versions available 2-5

space character, defined Glossary-5

specimen
- adding 6-11
- adding sample data manually 6-14 to 6-21
- adding to the project 6-11 to 6-12
- consensus sequence, defined Glossary-5
- creating automatically 6-12
- cumulative QV scoring in reports 10-13
- defined Glossary-5
- exporting 9-5
- removing sample data 6-21
- report, defined Glossary-5
- view, defined Glossary-6
- view, displaying 7-10

specimen (consensus)
- quality value, defined Glossary-5
- score, defined Glossary-5

Specimen Statistics Report 7-32, 10-13

starting the software 2-10

summary sequence, defined Glossary-6

system requirements, minimum 2-3

T

tab-delimited text file E-2 to E-4
  importing NT variant 4-32
technical support
  Services and Support Web site iii-xix
technical support, for computers with altered configuration G-1
template, saving 5-5
toolbar
  main 2-24
  viewing 2-25
translation tables
  amino acid abbreviations C-5
  complements C-3
  IUPAC diagrams C-3
  IUPAC/IUB codes C-2
  Universal Genetic Code C-4

U

uninstallation procedure 2-9
Universal Genetic Code C-4
upgrading
  from 1.0, 1.1, or 2.0 2-8
  software 2-8
Use Mixed Base Identification check box 3-7
user
  creating new 2-13
  information, changing 2-19
  login process 2-10
  new, login procedure 2-21
  privileges D-1
using
  aligned FASTA files 4-20
  Clear Range widget 8-17

V

Variant Styles tab, described 4-39
variants
  adding to projects 6-29
  aligned E-2 to E-4
  assigning styles 4-39
  changing unknown to known 6-25, 6-27
  creating text files E-2
  editing data 8-20
  importing 6-28
  importing into a project 6-32 to 6-33
  incorporating into project RDG 6-25
  promoting unknown to known 6-25
  viewing data 7-20 to 7-21, 8-20
versions of SeqScape software available 2-5
viewing
  project results 7-25
  toolbar 2-25

W

warranty
  claims G-2
  exceptions G-2
  for computers with altered configuration G-1
  limitations G-3
  period G-1
  rights and responsibilities 2-2
warranty period, effective date G-2
wizard, creating RDG 4-6
workflow
  analysis, display 3-2
  analyzing project 6-2
  creating project 6-2
  creating project template 5-2
  creating RDG 4-2
  editing data 8-2
  for viewing results 7-2
  reanalyzing data 8-2
  typical 2-28

Headquarters
850 Lincoln Centre Drive
Foster City, CA 94404 USA
Phone: +1 650.638.5800
Toll Free (In North America): +1 800.345.5224
Fax: +1 650.638.5884

Worldwide Sales and Support
Applied Biosystems vast distribution and service network, composed of highly trained support and applications personnel, reaches 150 countries on six continents. For sales office locations and technical support, please call our local office or refer to our Web site at www.appliedbiosystems.com.

www.appliedbiosystems.com

Applied Biosystems

Applera Corporation is committed to providing the world's leading technology and information for life scientists. Applera Corporation consists of the Applied Biosystems and Celera Genomics businesses.

Printed in the USA, 9/2003
Part Number 4346367 Rev. B

an **Applera** business